REPORT RESUMES

ED 003 520 24

ANALYSIS OF RESEARCH ON SELECTED ASPECTS OF EVALUATION IN ADULT EDUCATION.

BY- SUTTON, ELIZABETH W.

FLORIDA ST. UNIV., TALLAHASEE

REPORT NUMBER CRP-S-238 PUB DATE 66

REPORT NUMBER BR-5-8268

EDRS PRICE MF-$0.54 HC-$13.56 339P.

DESCRIPTORS- *ADULT EDUCATION, *RESEARCH METHODOLOGY, *EVALUATION TECHNIQUES, PROGRAM EVALUATION, *COMPARATIVE ANALYSIS, *EDUCATIONAL RESEARCH, METHODS RESEARCH, TALLAHASSEE, FLORIDA

AN EVALUATION OF ADULT EDUCATION RESEARCH ON SPECIFIC FUNCTIONS CONDUCTED AT CERTAIN INSTITUTIONS WAS REPORTED. IDENTIFICATION AND ANALYSIS WERE CONDUCTED FOR OCCUPATIONAL, LIBERAL, RELATIONAL, AND LITERACY EDUCATION FUNCTIONS. RESEARCH PROCEDURES WERE (1) TO IDENTIFY, SELECT, AND OBTAIN RESEARCH STUDIES, (2) TO MAKE CRITICAL ANALYSES AND EVALUATIONS OF ACCEPTED RESEARCH, AND (3) TO SUMMARIZE AND REPORT FINDINGS. ANALYSES AND EVALUATIONS WERE COMPLETED ON 40 RESEARCH REPORTS BY MEANS OF (1) DESCRIPTIVE STUDIES OF SINGLE PROGRAMS, (2) COMPARATIVE STUDIES OF TWO OR MORE PROGRAMS, AND (3) COMPARATIVE STUDIES OF THE OUTCOMES OF THE SAME PROGRAM APPLIED TO TWO OR MORE POPULATIONS. A TEAM APPROACH USING SCHOLARS FROM VARIOUS DISCIPLINES WAS SUGGESTED. (RS)

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

ANALYSIS OF RESEARCH ON SELECTED ASPECTS OF EVALUATION IN ADULT EDUCATION

Cooperative Research Project No. S-449-64

Elizabeth W. Sutton

Florida State University

Tallahassee, Florida

1966

The research reported herein was supported by the Cooperative Research Program of the Office of Education, United States Department of Health, Education, and Welfare.

ED 003520

ACKNOWLEDGMENTS

The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare under the provisions of the Cooperative Research Program. The writer expresses her gratitude to the United States Office of Education for the grant, and to Florida State University for its support of the contract. Appreciation is expressed to the university professors in the several institutions who provided copies of research reports, and to the personnel of the Robert Manning Strozier Library who obtained many theses and dissertations from other libraries.

The writer expresses sincere gratitude to all members of the Advisory Committee: To Dr. I. Douglas Brown, Dr. Herman L. Frick, Dr. Mildred Swearingen, for their counsel and encouragement; to Dr. George F. Aker, for assistance in identifying and in obtaining the research reports and for many other valuable suggestions; and finally, to the Chairman of the Committee, Dr. H. A. Curtis, for the guidance given during the conduct of the research and during the preparation of this report.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

Chapter

I. INTRODUCTION 1

Significance of Study
Limitations of Study
Methodology

II. OCCUPATIONAL STUDIES BY INSTITUTIONS . . . 23

Cooperative Extension
University Extension
Evening Colleges
Junior or Community Colleges
Public School Adult Education

III. LIBERAL STUDIES BY INSTITUTIONS 211

University Extension
Public School Adult Education
Great Books Foundation

IV. RELATIONAL STUDIES BY INSTITUTIONS 287

University Extension

V. SUMMARY AND CONCLUSIONS 293

BIBLIOGRAPHY 305

CHAPTER I

INTRODUCTION

The purpose of this study was to evaluate the quality of reported research dealing with the evaluation of adult education. The objectives were to identify, critically analyze, and evaluate the reported research which has been conducted under the auspices of specified institutions and which bears primarily upon specified functions. The specified institutions are: Cooperative Extension Service, University Extension, Evening Colleges, Community or Junior Colleges, Public Schools, and the Great Books Foundation. The specified functions are: Occuational Education, Liberal Education, Relational Education, and Literacy Education. These functions are defined as follows:

> Occupational Education: Those programs designed to improve the competencies of adults in vocational, technical, and professional endeavors.
>
> Liberal Education: Those programs designed to enhance the individual's satisfaction with himself and to enrich his cultural life. This function includes such program areas as the humanities--literature, the arts, history, religion, and philosophy--and education for wholesome recreation.

Relational Education: Those programs for which objectives are oriented to the development of a deeper understanding of the factors affecting human relationships and the improvement of one's adjustment in his personal contacts of life.

Literacy Education: Those programs designed to give adults the fundamental skills of literacy--basic reading competencies, fundamental number concepts, and speaking and writing proficiencies comparable to an elementary education. This function also includes those programs designed to improve competencies in the communication skills at any educational level.

As used in this investigation, evaluation research refers to research that is designed to determine the extent of those changes in the behavior of participants which are related to the objectives and which may be considered treatment effects of the given program. It also refers to research designed to determine the effectiveness of a method or to compare two or more methods employed for disseminating a given body of knowledge. Thus, evaluation research determines the extent to which objectives of a given program are achieved by the participants.

Significance of Study

With the rapidly expanding growth of adult education in the United States, adult educators have become increasingly concerned with the effectiveness of adult education. They have come to realize that vigorous evaluation is essential for the improvement of adult education programs. Among those who have expressed a concern about the quality of evaluation is Klineberg. He affirmed that evaluation should be restricted to fact-finding methods which yield evidence that is "objective, systematic and comprehensive," and that it should be distinguished from other forms of assessment which do not meet these criteria.[1]

However, there are many problems involved in conducting research to evaluate adult education. The problems stem from the fact that two mutually dependent factors operate. First, adult education is a field which is heterogeneous in its programs, practices, and interests, and second, adult educators hold varying concepts as to the purposes and objectives of adult education. Miller identified three positions in terms of differences of functions of adult education:

1. The intellectual orientation of the individual. Proponents of this position perceive the functions of adult education as helping people learn to see themselves more clearly, to think more critically, to cope with problems more intelligently and creatively, and to help aging adults to readjust their views of a changing world.

2. Increased competence in social role behavior. Proponents of this position see the function of

---

[1] Otto Klineberg, "The Problem of Evaluation," International Social Science Bulletin, Vol. 7 (1955), p. 347.

adult education in terms of the basic needs of individuals, in "becoming" and maintaining themselves as competent persons in their society.

3. Responsible functioning in the community and in society. Proponents of this position, designated as the modal position, perceive the function of adult education as helping to clarify the meaning of democracy, and to develop an understanding of its problems.[1]

The problems of evaluation in adult education are further compounded by the fact that many of the institutions and individuals in the field hold only vague objectives of any kind. Kreitlow, in writing on this point, observed that "this lack of definitiveness produces two key problems for research: what are the objectives of adult education and what are the outcomes?"[2]

Reports of research evaluating adult education are not only extremely difficult to identify in the literature but such reports are not readily accessible. Many studies have been conducted by persons engaged in various disciplines and in numerous publicly-supported service programs. Consequently, studies may not be reported in the literature to which adult educators usually refer. Furthermore, a substantial portion of the evaluation studies have not been published. These include theses, dissertations, and non-degree research conducted or directed by university professors in various disciplines or by practitioners in a multiplicity of agencies and organizations. To obtain such unpublished reports, the individual is obliged either to utilize the services of an inter-library loan system, to purchase a micro-film copy, or to contact the professor or agency directly responsible for the study. The following illustrations substantiate these statements.

---

[1] Harry L. Miller, "Philosophy and Issues," Review of Educational Research, Vol. 24 (1959), pp. 230-238.

[2] Burton W. Kreitlow, "Research in Adult Education," Handbook of Adult Education in the United States, ed. Malcolm S. Knowles (Chicago: Adult Education Association of the U.S.A., 1960), p. 114.

Of the studies reported by Wilder, four evaluated parent education and had been conducted either by State Departments of Mental Health or by Mental Health Societies.[1] As indicated in footnote references, one study was reported in the American Journal of Psychiatry, and another study had appeared in the American Journal of Public Health. Of the remaining two studies, one was available in the offices of the Louisiana Mental Health Society (New Orleans); the other from the Michigan State Department of Mental Health

---

[1] David E. Wilder, "Problems of Evaluation Research," An Overview of Adult Education Research by Edmund deS. Brunner et al. (Chicago: Adult Education Association of the U. S. A., 1959), pp. 243-273.

(Lansing). Two studies had been conducted by the American Cancer Society and bibliographic references indicated that one study was available from the Office in New York; the other from the Branch Office in Philadelphia. Of the evaluation studies on race relations programs, one was a doctoral dissertation and another was reported in the Journal of Abnormal and Social Psychology.[1]

The annual summary reports of "Research and Investigations in Adult Education" give the sources from which additional information or the completed reports may be obtained.[2] In most cases this information has been listed as follows: (1) "Inquiries to . . . ," followed by the name and address of the investigator, (2) "Available on loan from: . . . " followed by the name of the University Library or Department in which the research was conducted, and (3) "Micro-film available from: . . . " followed by the appropriate address. These reports of adult education research did not utilize evaluation as a topical heading until 1963; however, evaluation studies appeared under other organizational topics.

---

[1] The writer obtained the dissertation and learned that it had been conducted under the auspices of the National Conference of Christians and Jews.

[2] Adult Education Association of the U. S. A., "Research and Investigations in Adult Education," Adult Education, Vols. 5 (1955), pp. 114-127, 240-246; 6 (1956), pp. 234-243; 7 (1957), pp. 196-207; 8 (1958), pp. 195-207; 9 (1959), pp. 198-210; 10 (1960), pp. 195-218; 11 (1961), pp. 195-222; 12 (1962), pp. 195-216; 14 (1964), pp. 195-225; and 15 (1965), pp. 195-234.

It is evident that reports of research evaluating adult education are widely scattered throughout the literature and are not readily accessible. Hence, students of evaluation have been unable to study them comprehensively and systematically, a logical first step in the improvement of evaluation itself.

The only review of evaluation research in the literature to date which presents an analysis of evaluation research is that by Wilder.[1] In fact, his analysis of selected studies is of inestimable value to individuals interested in evaluation research. After specifying that his purpose was to discuss evaluation research in adult education, he stated:

> This chapter can do no more than describe a few representative studies, indicate the problems they met, their contributions and difficulties, and in conclusion suggest those considerations which must enter into any effective evaluation study.[2]

He chose and described selected studies which illustrated points that should be taken into consideration in evaluation research. These studies, conceptualized in terms of the experimental model, were discussed in terms of conceptualization, techniques and instruments, design, and analysis and interpretation.

---

[1] Wilder, op. cit.

[2] Ibid., p. 244.

Thus, it is obvious that the knowledge available on research evaluating adult education is spotty, scattered, and disorganized. No effort has been made to systematize the work over the full extent and the full range of the quality of evaluation research in adult education. Hence, the adult educator does not have a complete and systematic body of knowledge from which he can draw to reach objectively-based decisions regarding procedures for determining the effectiveness of adult education.

While there is a professional need for a comprehensive treatment of all evaluation research in adult education, this study did not attempt a task of such magnitude. It attempted rather to survey a defined segment of the total field, select and analyze studies representative of the most promising efforts, and to appraise the whole in terms of extent and quality.

Limitations of Study

This study was limited to the adult education research which evaluated either an adult education program or a method(s) employed for disseminating information. It was limited to that research conducted under the auspices of six specified institutions and which bears primarily upon four specified functions. The specified institutions were: Cooperative Extension Service, University Extension, Evening Colleges, Junior or Community Colleges, Public Schools, and the Great Books Foundation. The specified functions were: Occupational Education, Liberal Education, Relational Education, and Literacy Education. The study was further limited to the reported research which had been conducted in the United States during the period, 1950 to 1963, inclusive, and for which there was (1) a clear statement of purpose or problem, (2) description of method(s) employed and of population used, and (3) statement of results obtained and of conclusions drawn.

This study did not purport to cover the vast amount of research which appears in the more recently published literature of adult education. The investigator is cognizant of the fact that adult educators have directed considerable research efforts to such areas as adult learning, including psychological and physiological changes, interests, motivations, values, and attitudes; to adult education clientele, including their characteristics, extent of participation and cause of dropout; to the development of critera by which to improve programming for adults, including both planning and conducting adult education programs and the university graduate programs for preservice preparation of professional adult educators; and to the total problem of socialization of adults.

Moreover, this investigation did not include the usual professional type of inservice preparation programs conducted by General University Extension and Public Schools. Thus, programs designed specifically for continuing professional training in such fields as teacher-training, law, and medicine were excluded.

Two other types of decisions also may have limited this study. Reports of research were sampled in those areas in which more than ten investigations were identified. This limitation applied only to the evaluation research conducted under the auspices of the Cooperative Extension Service. Also, in reviewing the research of the Cooperative Extension Service, the investigator eliminated those studies which evaluated only a news-story, a circular, or a bulletin, per se. However, when considered a method in those studies which compared the effectiveness of two or more methods, the news-story, circular, and bulletin were included.

Methodology

The methods of research which have been employed during the conduct of this study may be described under three headings, namely, those used to identify, select, and to obtain the research studies; those used in the critical analyses and evaluation of each piece of accepted research; and those used in summarizing and in reporting the findings.

Procedures followed to identify, to select, and to obtain the evaluation research studies

To identify the evaluation research studies, the investigator followed a carefully-outlined plan which included first the listing of standard bibliographic indexes, reviews of research, and other references judged to be critical sources for the development of the bibliography of studies. Before making a detailed study of any bibliographic index or reference source, each one was examined carefully in order to develop a list of subject entries which were especially appropriate for investigating that particular source. The lists of subject entries developed included the two which were used to examine the Bibliographic Index, the fourteen for the 1960 edition of the Encyclopedia of Educational Research, the twenty for Education Index, and the twenty-five for Sociological Abstracts. All volumes of these indexes were investigated to cover the specified time interval, 1950-1963, inclusive. Other standard bibliographic sources which were examined similarly included Psychological Abstracts, Review of Educational Research, Index to Doctoral Dissertations, Dissertation Abstracts, and Master's Theses in Education.

During the process of this phase of investigation other bibliographic sources were added to the list. Many were specially prepared bibliographies which were obtained by directing a letter of request to the compilers or to the appropriate agency. Other sources were professional books which included two published by the Adult Education Association and two written by eminent sociologists in which the findings of adoption-diffusion research were synthesized.[1] Finally, a total of approximately fifty bibliographic sources were investigated, all of which contained several studies which appeared to be those for which the investigator was searching. However, the listed title was often not an adequate basis upon which to determine the appropriateness of the study. Those bibliographic sources which proved to be of most value were the ones which contained either summary reports, a procedural outline (purpose, method, and findings), or abstracts of studies. Among those sources were: (1) Review of Extension Research, including fifteen different publications, (2) "Research in Cooperative Extension Work," a series of manuscripts prepared annually since 1957 by the College of Agriculture of the University of Wisconsin, (3) Summaries of Studies in Agricultural Education, a series of bulletins published annually by the United States Department of Health, Education, and Welfare, (4) "Research and Investigations in Adult Education," a report of research which has appeared annually since 1955 in Adult Education, and (5) Dissertation Abstracts.[1]

---

[1] Edmund DeS. Brunner et al., An Overview of Adult Education Research (Chicago: Adult Education Association of the U.S.A., 1959); Malcolm S. Knowles (ed.), Handbook of Adult Education in the U.S.A. (Chicago: Adult Education Association of the U.S.A., 1960); Herbert F. Lionberger, Adoption of New Ideas and Practices (Ames, Iowa: The Iowa State University Press, 1960); and Everett M. Rogers, Diffusion of Innovations (New York: The Free Press of Glencoe, 1964).

✓

Originally, approximately 250 studies were listed in the investigator's bibliography of the research conducted by the Cooperative Extension Service to assess the effectiveness of programs and methods as specified previously. As a result of efforts to develop a set of criteria by which a representative sample could be drawn, the investigator discovered three types of ambiguities: (1) the same study had been listed in bibliographic sources under different titles and by different authors, (2) additional reports or publications of the same study were found, and (3) the studies on adoption of farm and home practices could not be differentiated precisely from studies which related to diffusion of agricultural innovations.

---

[1] With the exception of Dissertation Abstracts, a complete bibliographic entry for these five references is included in the first section of the Bibliography, pp. 305-306. In the event the primary summary report found for a study was in the Dissertation Abstracts, the appropriate citation follows the listing of the study in the Bibliography. For the most part summary reports as used in subsequent sections of this report refers to the five bibliographic sources as listed.

The foregoing difficulties necessitated the expenditure of considerable time and effort in the further investigation of each listed study, and the decision was reached that no discrete categories could be developed under which all Cooperative Extension Studies could be subsumed unless the investigator obtained and read each of the studies. Finally, duplicated studies were eliminated, and the Bibliography which contains a list of 202 studies conducted under the auspices of the Cooperative Extension Service was organized into three sections: (1) "Behavioral Change," including those studies (61) which were conducted primarily to determine the extent to which participants changed in attitudes, skills, or practices, (2) "Effectiveness of Method(s)," including those studies (43) which were conducted primarily to determine the effectiveness of a method or to compare two or more methods for disseminating information, and (3) "The Adoption-Diffusion Studies," consisting of those studies (98) conducted primarily to ascertain kinds of factors and types of communication channels which affect the acceptance and diffusion of agricultural (both farm and home) practices. The actual process of classifying studies under three sections listed involved personal judgment at many points.

For analysis, a sample of twenty-two studies was selected on a judgmental basis from the 202 studies which had been conducted under the auspices of the Cooperative Extension Service. The investigator is aware that a sample so selected required personal judgment and may be somewhat biased. However, careful consideration was given to the following criteria in the selection of this sample:

1. Studies which appeared to be representative of the best research procedures, specifically studies for which the research design had been carefully developed, and for which data had been obtained and analyzed in an acceptable manner.

2. Studies from each of the three sections listed above, with priority given to studies listed in Sections A and B, "Behavioral Change," and "Effectiveness of Method(s)," respectively.

3. Studies representative of those conducted or sponsored by well-known universities, including established Cooperative Extension Service and Agricultural Experiment Stations.

4. Studies which evaluated various types of Cooperative Extension Service Programs.

The studies included in the third section, "The Adoption-Diffusion Studies," were not conducted primarily to evaluate a program but since they determined the extent of adoption of selected practice(s), they contain evaluation data. Several otherwise valuable studies which fell in this section of the Bibliography were deleted because the investigator found no reference made to determining the extent of adoption. A number of the studies listed under the Cooperative Extension Service were so included on the basis of their having been reported in reliable sources as research and upon the best judgment the writer could form from information given in bibliographic sources, but not on the basis of having read each and every one.

Reports of studies were obtained from four different sources. Several studies were reported in professional journals and periodicals, all of which were available in the University Library. Dissertations and theses were borrowed through the Inter-Library Loan System or procured on micro-film. Many studies were obtained in response to individual letters of request directed to appropriate personnel and agencies. The correspondence necessary to obtain these research reports and to acknowledge the same amounted to more than 100 letters. Only six letters requesting copies of studies were not acknowledged. Twelve studies which appeared to be those of evaluation research were not available. According to the several respondents either the studies had not been completed or no report of the study could be located by the respondent.

Thus, in the process of identifying, selecting, and obtaining the research studies the investigator examined thousands of titles. Most studies were rejected during the initial investigation and others were eliminated upon further investigation of summary reports and research reviews. Finally, of the original bibliography of 470 studies, 250 were rejected. Reasons for rejecting these studies as revealed by the tabulations made on the basis of the first specified limitation which was discovered may be summarized as follows:

1.	Misleading titles of studies	126
2.	Conducted by institutions other than those specified	60
3.	Did not meet the research criteria	32
4.	Subjects were regularly enrolled college or university students and the investigation had been conducted by instructors for a given course	23
5.	Studies conducted prior to the specified date, 1950	7

6. Not related to one of the four specified functions of adult education 2

Accordingly, 220 evaluation studies were identified and classified by functions over institutions. The extent of evaluation research which has been conducted during the period, 1950 to 1963, inclusive, is depicted in Table 1. The 220 evaluation research studies are listed by institutions in the Bibliography.

TABLE 1

NUMBER OF EVALUATION STUDIES IDENTIFIED--CLASSIFIED BY FUNCTIONS AND BY INSTITUTIONS

Functions	Coop. Extension Service	Univ. Extension	Evening Colleges	Jr. or Community College	Public School Adult Educ.	Great Books
Occupational	202	5	2	1	2	...
Liberal	...	1	...	...	3	3
Relational	...	1	...	...	...	...
Literacy	...	...	...	...	...	...

Methods employed to analyze and evaluate the research

The investigator read and studied each piece of accepted research in a meticulous manner after which a digest of the study was written. In the writing of each digest special effort was made to interpret accurately and to describe succinctly each study in terms of the following elements of research:

1. The stated purpose, objectives or problem and hypotheses;
2. The design of the study;
3. The population for the study and the method employed in the selection of the sample;
4. The criterion measurement which was used including the development of instruments and the efforts made to establish their reliability and validity;
5. The procedures used in the analysis of data;
6. The level of confidence selected or accepted to designate statistical significance if the data were analyzed statistically;
7. The manner in which findings were reported and whether or not tables, graphs, or figures were used to present findings and supporting data;
8. A summary of the results of the study; and
9. A summary of the conclusions which were drawn.

After writing the digest of the study the investigator made an evaluation of it. First, this evaluation was made in terms of the technical qualities of research, namely, appropriateness of method, adequacy of sample, type and quality of measures, confidence level of reported results, and soundness of conclusions. Secondly, the evaluation considered each study in terms of its importance as a contribution to practice, or as a contribution to knowledge. A study was considered to contribute to practice if it dealt with a technique or a body of subject matter that is in the common domain, or if it refined or improved practice or invalidated existing practices. A study was considered to contribute to knowledge if it tested or extended existing theory or if it opened a new body of theory.

Methods employed to summarize and report findings

To summarize and report findings the investigator has dealt with the identified evaluation research studies pertaining to the occupational, liberal, and relational functions of adult education, respectively. No studies relating to the literacy function of adult education were identified. For each of the three functions, the research conducted under the auspices of each of the appropriate institutions has been considered. Because of differences in the quality of studies, the investigator either wrote a critical analysis of each study or formulated categories in which studies of similar research methodology were classified. In the latter case critical analyses were made of selected studies and similarities and differences between the reported study and the corresponding remaining studies were noted. In dealing with the selected sample of twenty-two studies conducted under the auspices of the Cooperative Extension Service, the investigator has cited other studies listed in the Bibliography which appear to have the same research characteristics as those classified in a given category. The organizational pattern is described in more detail at appropriate points in the report.

A summary evaluation was made of the studies classified in each category or of the studies conducted by a given institution. This evaluation was made in terms of the technical qualities of research and in terms of the contributions made to practice or to knowledge as specified previously. Finally, a summary statement of the present status of evaluation research is offered and a brief set of recommendations is made for the improvement of the practice of evaluation in the field of Adult Education. The results of this work are reported sequentially in the sections immediately following.

CHAPTER II

OCCUPATIONAL STUDIES BY INSTITUTIONS

The purpose of this chapter is to report the analysis and evaluation of selected research reports which are related primarily to the occupational function of Adult Education. A substantial amount of research effort has been directed toward evaluation of programs which were designed to improve the competencies of adults in vocational endeavors. Table 1[1] clearly indicates the extent of this research and reveals that the number of identified evaluation studies which has been conducted under the auspices of the specified institutions is as follows:

1.	Cooperative Extension Service . . .	202
2.	University Extension	5
3.	Evening Colleges	2
4.	Junior or Community Colleges	1
5.	Public School Adult Education . . .	2.

The studies for each institution are reported separately. First, a critical analysis is presented of the studies, after which a summary evaluation of the group of studies or of a sub-group of studies is made in terms of the technical qualities of research and in terms of contribution made either to practice or to knowledge.

---

[1] *Supra*, p. 18.

Cooperative Extension Service

An analysis of the twenty-two studies selected as a sample of those conducted under the auspices of the Cooperative Extension Service reveals a wide range of studies both in terms of purposes or problems and of quality. In terms of purposes or problems these studies range from those in which it was the investigator's purpose to secure opinions about a program or a method to those in which it was the investigator's purpose to formulate and test a theory experimentally. In terms of quality they range from descriptive status-surveys with data reported in frequencies and percentages to those with carefully controlled experimental designs and quite sophisticated analytical data treatments.

The sample of Cooperative Extension Service studies was classified in three major categories. Category I consists of those studies which related selected personal, social, or situational factors to the adoption or nonadoption of recommended farm and home practices. As pointed out previously, the studies in this category were not conducted primarily to evaluate Cooperative Extension Programs, but inasmuch as they determined the extent of adoption of recommended practices, they contain evaluation data. This category has been designated "Validity of Predictors of Adopters."

The studies embodied in Category II are those which were designed primarily to determine sources of information of most value to the farmer or homemaker at stated stages or times in the adoption-diffusion process. Likewise, these studies were not conducted primarily for evaluation purposes, but contain evaluation data relevant to the extent of adoption of a recommended practice or practices at specified stages or times. This category has been designated "Informational Sources: Stated Stages of Adoption."

Category III, the major category, has been designated as "Methods," using the word in its generic sense. This category has been subdivided into four sub-categories as follows:

1. III-A. Those studies in which one treatment was applied and the results were assessed in relation to the participants as an undifferentiated group.
2. III-B. Those studies in which one treatment was applied and the results were assessed in relation to differentiated groups of participants.

3. III-C. Those studies in which two or more treatments were applied and the results were assessed in relation to the participants as an undifferentiated group.
4. III-D. Those studies in which two or more treatments were applied and the results were assessed in relation to differentiated groups of participants.

For each category, the studies were analyzed, after which an evaluation of the group of studies was made. This evaluation was made in terms of the technical qualities of research as specified previously and in terms of the contributions made either to practice or to knowledge. The results of this work are reported sequentially in the sections immediately following.

Category I: Validity of Predictors of Adopters

Three of the investigations in the selected sample of twenty-two studies were classified in this category. The primary purpose of each study was to determine whether selected factors, either personal, social, or situational, were associated positively with adoption of recommended farm practices. Each researcher used the concept of adoption as the criterion variable. Because of the differences in purpose, design, and statistical treatment, and because each is illustrative of a substantial amount of research which has been conducted in recent years under the auspices of the Cooperative Extension Service, a report of each study is presented.

Brandner formulated and tested the theory that an innovation that is congruent or consonant with a previous, favorably evaluated practice will be adopted more rapidly than will one which is not congruent with a previously adopted practice.[1] He termed this theory "evaluation of congruity." In conducting the research Brandner made three kinds of investigations.

First, through a comprehensive review of the literature, he identified more than fifty factors found to be associated either positively or negatively with adoption. He pointed out that when such factors were identified as attributes or as nonattributes of Kansas farmers, one could reliably predict those farmers who would adopt hybrid sorghums more rapidly and those who would adopt hybrid sorghums less rapidly.

Secondly, a reanalysis of selected research studies was made to test the applicability of the theory of "evaluation for congruity" to the data of each study. The evidence from the reanalysis of data drawn from the fields of sociology, anthropology, psychology, and medical science substantiated congruence as an explanatory principle, and supported the theory that this factor may be more powerful in adoption than are those previously identified.

---

[1] Edward Lowell Brandner, "Evaluation for Congruence as a Factor in Accelerated Adoption of an Agricultural Innovation" (unpublished Ph.D. dissertation, University of Wisconsin, 1960).

Finally, a primary study was conducted to test the theory. The locale selected for the study was an ideal one in that the growing of hybrid sorghums, the recommended practice, was a recent innovation adapted to all areas of Kansas. Production records showed sorghums to be of much more importance economically in the southwestern areas than in the northeastern areas of Kansas. Conversely, hybrid corn had been an economically important crop in the northeastern areas but not in the southwestern areas. Thus, the ecology of the state provided the control needed to determine if those farmers who previously had evaluated hybrid corn favorably were influenced by that experience to adopt hybrid sorghums more rapidly than were those farmers who lacked experience with hybrid corn.

A cluster sample of farmers was drawn from one county in northeastern Kansas which represented the commercial corn area of the state. A second cluster sample of farmers was drawn from a county in southwestern Kansas which represented the non-corn area of the state. Both areas produced grain sorghums. Data were obtained by personal interviews using a carefully-constructed and pretested schedule. This schedule included standard questions pertaining to the identified factors, a projective device, open-end questions and probes. Open-end questions, followed by probes were used in an attempt to obtain evidence of underlying factor(s) that might have influenced farmers to adopt hybrid sorghums. To illustrate: Each respondent was asked two open-end questions sequent to the question, "How many acres did you have in hybrid milos last year?" The subsequent open-end questions were, "Weren't you afraid the new hybrid milos wouldn't do as well as the regular milos?" and "What gave you the idea or made you think hybrids would be better?" Probes with the two questions were, "Why?" . . . "Anything else?" . . . "Were there other reasons?" punctuated with pauses and silences. Therefore, any mention of experience with hybrid corn as a reason for trying hybrid sorghums came as a spontaneous response from the interviewee. The projective device was a photo used in an attempt to discover the respondent's reference groups. It also stimulated interest, helped to put the respondent at ease, and provided a break in the interview.

Brandner interviewed 309 rural residents in the sample areas in his search for all the farmers who had tried hybrid sorghums before 1958 and for their nearest neighbors. Eighty-six schedules were completed. Farmers were categorized into northeastern adopters (25 in number), and non-adopters (26); and into southwestern adopters (18), and non-adopters (17) in order to make comparisons between categories and between sections of the state. According to state-wide data obtained, this sample was representative of those who had, and had not, grown hybrid corn.

All data were analyzed statistically. Differences were tested for significance by chi square and by the t-test, and nothing less than the .05 level of confidence was accepted. The report contained a lengthy interpretation of findings. Data were presented in forty-two tables interspersed within the context. One table, consisting of three pages, offered a summary of factors that either previous research or logic had associated with acceptance or rejection of technological innovations and showed results of comparisons for each factor between designated areas and categories. The findings which appear to be of most significance are summarized briefly in the following paragraphs.

Six factors indicated that northeastern farmers would adopt hybrid sorghums more rapidly than would the southwestern farmers. However, only three of the six factors were statistically significant. Against these six factors, twenty-three factors indicated that the southwestern farmers would adopt hybrid sorghums more rapidly than would the northeastern farmers. Nine of the twenty-three were statistically significant. Thus, the bases for adoption were three to one in favor of the southwest farmers.

Against this overwhelming prediction based on factors found to be associated with adoption from previous research, the farmers in the northeastern area adopted hybrid sorghums four times as rapidly as did those farmers in the southwestern area. This indicated that the influence of hybrid corn in the Northeast outweighed nine statistically significant differences favoring the southwest. The data revealed that the chances were better than 999 in 1,000 that the idea that hybrid sorghums were like hybrid corn resulted in hybrid sorghums being adopted much more rapidly by northeastern farmers than by southwestern farmers.

Brandner drew the following as tentative conclusions and specified that they should be tested by further research:

1. That one of the important evaluative processes an individual follows in deciding whether to accept

or reject a given innovation is to examine the practice in terms of whether it appears to be congruous or consonant with some other practice on which a previous favorable evaluation has been made.
2. That each adoption situation may be characterized by greater or lesser reliance on this process of "evaluation for congruity."
3. That an individual will accept an innovation faster when he sees the innovation as congruous or consonant with previous favorably evaluated practices.[1]

Marsh and Coleman examined the relationship of farmer characteristics to the adoption of individual farm practices, and the consistency with which the characteristics differentiated between adopters and nonadopters of sixteen specific practices.[2] Two questions guided their research:

1. What socioeconomic characteristics of farmers are related to their adoption of recommended farm practices?

2. Are there characteristics related to the adoption of all or most practices, or do the characteristics that are related to adoption vary widely by practice?

Data were obtained by personal interviews with the 393 farm operators residing in thirteen different neighborhoods of one Kentucky county. These neighborhoods were selected on the basis of the researchers' judgment that they represented all areas of the county. Sixteen farm practices were selected and the relationships of twenty-one factors to the adoption of these practices were examined. The factors included status items, items reflecting exposure to potential sources of information, and items on social participation.

---

[1] Ibid., p. 154.

[2] C. Paul Marsh and A. Lee Coleman, "The Relation of Farmer Characteristics to the Adoption of Recommended Farm Practices," Rural Sociology, Vol. 20 (1955), pp. 289-296.

Most of the analyses were based on simple cross-tabulations of each factor by adoption or nonadoption of each of the sixteen practices. Chi square was used to test the null hypotheses implicit in the comparisons of 336 relationships (21 times 16 practices). An existing relationship was accepted only if chi square was significant at the .05 level of confidence. Results were clearly presented in five tables with footnotes indicating chi square and confidence level. Two of the five tables illustrated the way in which the 336 tables were set up.

The following factors were found to be associated with the adoption of most of the practices:

1. One factor (score on the Sewell Socioeconomic Status Scale) was related to the adoption of all sixteen practices.
2. Two factors (education, and having talked with an agricultural agency representative in the past two years) were each associated with fifteen of the practices.

3. Three factors (Farm Bureau participation, having read one or more farm bulletins in the past two years, and the value of crops and products sold) were each related to the adoption of fourteen practices.

4. Three factors (having attended one or more farm meetings sponsored by agricultural agencies in the county, size of farm operation, as measured by productive man-work units, and social participation, as measured by Chapin Scale scores) were each associated with the adoption of at least twelve of the sixteen practices.

The relationship of each of the three top-ranking factors (education, socioeconomic status, and contact with agricultural agency representatives) to the adoption of each of the sixteen practices was examined by successively holding constant one of the remaining two factors. Relationships found were somewhat less consistent:

1. Socioeconomic status was consistently related to adoption in two out of three control categories: among farm operators who had completed only eight grades or less of schooling, and among those who had talked with an agricultural agency representative. This factor was positively related to adoption of each of the fourteen principal practices.

2. Education was associated with adoption of only six practices among those farmers with a high socioeconomic status. However, among operators who had talked with an agricultural agency representative, education was significantly associated with the adoption of eleven practices.

3. Having talked with an agricultural agency representative was related to adoption of thirteen practices within the group with only eight grades or less of schooling; among those farmers of high

> socioeconomic status, this factor was related to the adoption of only eight practices; and, among those of lower status, it was related to adoption of only four practices.
>
> The investigators concluded that for the study:
>
> 1. On the basis of the simple cross-tabulations, a set of six factors discriminated between adopters and nonadopters of at least fourteen of the sixteen practices studied, and several other factors had fairly high discrimination. However, on the basis of the controlled analysis of three of the top-ranking factors, the relationships were less consistent.
> 2. While the data do not permit a definite conclusion, it is noteworthy that all of the relationships in the controlled analysis are in the same direction as in the simple cross-tabulations.[1]

Rogers employed conceptual variable analysis to determine how well technological change can be predicted by sociological factors and which sociological factors are the best predictors of technological change.[2] Drawing upon sociological research, Rogers set forth the theory that all concepts may be expressed either as discontinuous or continuous variables. For his purposes, he dealt only with concepts which could be expressed as continuous variables. These he operationally defined as conceptual variables and proceeded to develop a methodology for handling them as such. Thus, the conceptual variable, technological change, was the degree to which individuals accepted new technological practices. Rogers used this conceptual variable as the dependent variable. In addition, the following six independent conceptual variables were selected: (1) change orientation, (2) communication competence, (3) status achievement, (4) cohesion with the locality group, (5) family integration, and (6) cohesion with the family group. These six conceptual variables were derived from a situational analysis of technological change. Each was selected on the basis of five criteria which required each concept to be: (1) sociological, (2) manipulatable or dynamic, (3) general or abstract in nature, (4) operationalizable, and (5) expected to be highly related to the dependent variable

---

[1] Ibid., p. 296.

[2] Everett M. Rogers, "A Conceptual Variable Analysis of Technological Change," Rural Sociology, Vol. 23 (1958), pp. 136-145.

The next step in Rogers' variable analysis was to develop operational scales or indexes to measure each conceptual variable. An operation was defined as the empirical referent of a concept. For example, the score on the adoption of farm practice scale was the operational measure of the concept of technological change. Thus, a general hypothesis, the postulated relationship between two conceptual variables, was tested by testing an empirical hypothesis of the relationship between two operational measures. For example, one empirical hypothesis was "the adoption of farm practices scores varies directly with status achievement scores."[1]

Accordingly, Rogers carefully defined each conceptual variable and constructed an operational index or scale to measure it. On the basis of prior research he stated a general hypothesis postulating the relationship of each independent conceptual variable to technological change. An empirical hypothesis was formulated, as illustrated above, to test each corresponding general hypothesis.

Briefly, he postulated that a positive relationship was expected between adoption scores and (1) change orientation scores, (2) status achievement scores, (3) communication competence scores, and (4) extralocality orientation scores. He defined the concept of cohesion as the degree to which an individual accepts the roles prescribed by a reference group. The degree to which an individual fulfills the role expectations of the locality reference group (neighborhood and community) was operationally measured by the index of extralocality orientation; hence, a positive relationship was expected. A negative relationship was expected between adoption scores and (1) the kinship orientation scores, and (2) family integration scores. Each of the specially constructed indexes was tested for validity and reliability, and all were accepted as operational measures of the concepts.[1] Data were obtained by personal interviews with the 148 farm operators residing in an Iowa rural community in 1955.

---

[1] Ibid., p. 138.

The six general hypotheses were tested by applying the statistical technique of multiple correlation. The .01 and .05 levels of confidence were specified as operating levels. Results of these correlations were interpreted as follows.

The zero-order correlations of the independent variables with the dependent variable, technological change, revealed that all were in the expected direction as suggested by the general hypothesis, but only three were significantly different from zero. Neither extralocality orientation, family integration, nor kinship was significantly correlated with adoption. These variables, however, were retained in the multiple correlation analysis because some were related to some of the other independent variables.

---

[1] Rogers referred to his original report of this research for a discussion of these procedures: "A Conceptual Variable Analysis of Technological Change" (unpublished Ph.D. dissertation, Iowa State University, 1957).

To determine the relationship between the dependent variable and the combined effect of the six independent variables, the coefficient of multiple correlation was computed. The coefficient of multiple determination was .167 which meant that 16.7 per cent of the variance in the adoption scores could be accounted for by the combined effect of the six independent variables.

Fifth-order partial correlations were used to determine the relationship between each independent variable and the dependent variable while controlling the effect of the other five independent variables. Three variables, change orientation, communication competence, and status achievement were found to be significantly related to technological change while three, cohesion with locality group, family integration, and cohesion with kinship group, were not significantly related to technological change.

Essentially, the conclusions were the findings from the fifth-order partial correlations since the significance of the relationship of each of the first three independent variables was not destroyed by controlling for the combined effect of the other five independent variables. Moreover, since almost 17 per cent of the variance in the dependent conceptual variable was explained by the variation in the six independent variables, it was concluded that much of the variation in technological change remained to be explained by other than these six conceptual variables.

The research methodology for each of the three foregoing studies was appropriate for the purposes as stated. The studies conducted by Brandner and by Rogers represent the most sophisticated type of research identified in those studies which are included in Category I, and as will be shown later, in Category II, also. These studies contribute principally to knowledge rather than to practice.

The samples were adequate for stated purposes. Results from the study by Marsh and Coleman may be generalized only to the farm operators in the specified county. Findings from the studies by Brandner and by Rogers add to the body of knowledge in general because the theory developed in each may be applied to other situations.

Data for all three studies were obtained from personal interviews. Brandner reported the pretesting of his schedule nine times, and Rogers established both validity and reliability of the scales and indexes used. Marsh and

Coleman gave no description or information on the interview schedule. However, it may be assumed that data were obtained in a reliable manner since the investigators are competent sociologists and are experienced in the field of adoption-diffusion research. Moreover, the report analyzed was probably somewhat condensed since it appeared as an article in a periodical.

Statistical procedures for analyzing data were sound and appropriate for stated purposes in all three studies. However, Marsh and Coleman applied chi square to data obtained from respondents of a sample which was not selected on a probability basis. They reported this fact but offered no justification for its use other than previously reporting that the thirteen neighborhoods were judged to be representative of all areas of the county. The .05 and higher levels of confidence were used for determining significance in all three cases. These levels are generally utilized in this type of research. In all cases, the conclusions as drawn appear to be sound and warranted.

The three studies classified in Category I, "Validity of Predictors of Adopters," appear to be characteristic of a substantial amount of the research listed in the

Bibliography.[1] However, based on a careful investigation of summary reports of the studies listed, it is the considered judgment of the investigator that the studies by Brandner and by Rogers were the only ones where major contributions were made to the advancement of knowledge. The major contribution of other studies was to an applied undertaking and such contributions as they may have made to the advancement of basic knowledge was a minor portion of their value. Other studies which are illustrative of this category include those by Bonser, Copp, Duncan, Fliegel, Havens, Kimball, Lionberger, Spencer, Wilkening, and Young.[2]

---

[1] *Infra*, pp. 320-332.

[2] H. J. Bonser, "Rural Community Clubs in Tennessee--Their Bearing on Adoption of Farm and Home Practices" (unpublished Ph.D. dissertation, Pennsylvania State University, 1957); James H. Copp, *Personal and Social Factors Associated with the Adoption of Recommended Farm Practices among Cattlemen*. Technical Bulletin 83 (Manhattan: Kansas Agricultural Experiment Station, 1956); James A Duncan, "The Relationship of Selected Cultural Characteristics to the Acceptance of Educational Programs and Practices among Certain Rural Neighborhoods in Wisconsin" (unpublished Ph.D. dissertation, University of Wisconsin, 1953); Frederick C. Fliegel, "A Multiple Correlation Analysis of Factors Associated with Adoption of Farm Practices" (unpublished Ph.D. dissertation, University of Wisconsin, 1955); A. Eugene Havens, "Social Psychological Factors Associated with the Differential Adoption of New Technologies by Milk Producers" (unpublished Ph.D. dissertation, Ohio State University, 1962); William J. Kimball, "The Relationship between Personal Values and the Adoption of Recommended Farm and Home Practices" (unpublished Ph.D. dissertation, University of Chicago, 1960);

The preceding section dealt with the three investigations from the selected sample of twenty-two studies which fell in Category I, "Validity of Predictors of Adopters." Each study has been analyzed and evaluated in terms of the technical qualities of research and in terms of its contribution to practice or to knowledge. Ten other studies listed in the Bibliography which are illustrative of the research classified in Category I were identified. The following section is devoted to those studies classified in Category II which has been designated as, "Informational Sources: Stated Stages of Adoption."

Category II: Informational Sources--Stated Stages of Adoption

The purpose of this section is to report the analysis and evaluation of the two studies in the sample of twenty-two investigations which fell in Category II. These two studies sought to identify those informational sources which

---

Herbert F. Lionberger, "Some Characteristics of Farm Operators as Sources of Farm Information in a Missouri Community," Rural Sociology, Vol. 18 (1953), pp. 327-338; George Elwood Spencer, "Value-Orientations and the Adoption of Farm Practices" (unpublished Ph.D. dissertation, Cornell University, 1958); Eugene A. Wilkening, "Informal Leaders and Innovators in Farm Practices," Rural Sociology, Vol. 17 (1952), pp. 272-275; and James N. Young, "The Influence of Neighborhood Norms on the Diffusion of Recommended Farm Practices" (unpublished Ph.D. dissertation, University of Kentucky, 1959).

were of most value to farm operators at stated stages or times in the adoption process. Studies classified in this category are similar to those in Category I, particularly in terms of the method employed to obtain data and in terms of the criterion measurement applied. Likewise, because these two studies are representative of a substantial amount of the research conducted under the auspices of the Cooperative Extension Service, both studies are reported.

Copp, Sill, and Brown examined the use of information sources by farm operators in the farm practice adoption process and developed a model which accounted for the function of information sources in this process.[1] Data were obtained by personal interviews with a random sample of 175 dairy farmers in a western Pennsylvania County. Each farmer recounted his experience with each of the three recommended practices in response to a series of guiding questions. The schedule also elicited responses relating to items concerning information source exposure and evaluation, production practices, farm mechanization, characteristics of the farm operating unit and characteristics of the farm operator.

---

[1] James H. Copp, Maurice L. Sill, and Emory J. Brown, "The Function of Information Sources in the Farm Practice Adoption Process," Rural Sociology, Vol. 23 (1958), pp. 146-157.

In the first analysis of the data, the farmer's experiences with the recommended practices were classifed into five arbitrary stages:

1. Awareness--the farmer first hears of the practice;
2. Interest--the farmer feels the practice is a workable solution for a farm problem;
3. Acceptance--the farmer feels the practice would be of value on his farm;
4. Trial--the farmer tries the practice; and
5. Adoption--the farmer continues to use the practice.

In the second analysis a comparison was made of the information-seeking behavior of farmers who were at different stages in the adoption process. The adoption progress score, one of the dependent variables used, was based on the latest stage attained by a farm operator in the adoption of a given practice. Informational sources were classified into eight categories: (1) magazines (farm magazines and papers), (2) radio, (3) printed extension (circular letters and bulletins), (4) oral extension (office calls, meetings, farm visits, method demonstration, tours, and field days), (5) peer influences (neighbors, friends, and relatives), (6) commercial media (both oral and printed), (7) the classroom (vocational agriculture and veterans' training), and (8) other, a residual category. Presented in percentages in a table and interpreted, these data revealed that there was not a one-to-one relationship between the citation of informational sources and stages in the adoption process. More than one source was mentioned within each of the earlier stages, and sometimes an operator passed through a stage without indicating other sources of information. However, the sources of greatest importance for each of the first four stages of adoption were: magazines, radio, and printed extension for the awareness stage; oral extension and peer influences for the interest stage; friends, neighbors, and relatives for the acceptance stage; and printed and oral extension sources for the trial stage. No informational sources were reported for the adoption stage because "the results of a successful trial were considered self-evident by the farm operator."[1]

The next step in the analysis was the comparison of the mean progress scores of farmers who cited different informational sources at a given stage. These data also were reported in tabular form and interpreted. Scores of average, above average, or below average for farmers who had cited sources of information for each stage were discussed. Summarizing, the investigators pointed out that there seemed to be evidence that farmers relying upon friends and neighbors for information in the early stages of the adoption process tended to have lower progress scores than did those who did not cite peer influences, and that there was no key informational source for a given stage.

---

[1] *Ibid.*, p. 151.

Lastly, to account for the differential use of sources of information as farmers progress toward adoption, a model was developed. A rationale was established for five considerations in its construction. These were: (1) institutionalization of informational sources, (2) the temporal sequence of informational sources, (3) the possibility of negative recommendations, (4) scheduling limitations of the sources, and (5) the need for local legitimation. When this model was applied to the data, the direction of the differences in the first three stages was consistent with the model's predictions in eight out of nine instances. The results from the sign test indicated that the probability of securing as good a fit between model and data through the operation of random factors was less than one in fifty. Even though the variation in progress scores for the trial stage was too limited to compare the means, it was noted that more than 80 per cent of the farmers made the expected selection of sources in this stage for each of the practices. When the mean progress scores for expected sources and for other sources of farm information for four stages in the adoption of the three recommended practices were compared, the results suggested that the farmers who selected sources in conformity with the model made the greatest progress toward the adoption of a practice. Moreover, when the differences of the means of progress scores, classified by presence or absence of peer informational sources in awareness and interest stages for the three practices, were tested, it was found in five of the six comparisons that farmers who mentioned other farmers as sources of information had significantly lower scores than did those who did not cite peer influence for these stages. Three of these comparisons were significant at the .05 level of confidence, and two were significant at the .001 level.

It was concluded that in a psychological sense, the basic problem for the farmer during the awareness and interest stages is building a cognitive structure of the practice. Furthermore, data in this study indicated that peer influences are less effective than are other influences in developing a cognitive structure which leads to the attainment of later stages in the adoption process. "The farmer who learns from his peers is learning second- or third-hand information, which may have lost much of its original accuracy."[1]

Rogers and Burdge reported a study which was originally designed to determine the role of the Ohio Agricultural Experiment Sub-station Farm in the adoption and diffusion of innovations by muck vegetable growers.[2] The purpose of the available report was to supplement an oral presentation which had been made previously and to report additional findings from a further analysis of the data.

A muck vegetable grower was defined as one who grew vegetables such as onions, celery, radishes, potatoes, carrots, spinach, parsley, or lettuce on muck soils. Applying this definition forty-nine muck vegetable farms in Ohio were identified in 1959. Three major reasons for the importance of this study were reported:

1. This was the first study to investigate the diffusion of innovations from a specific research facility to farmers. Because it had been found that there was a period of several years between the release of an innovation and its widespread adoption by farmers, an eventual goal of this study was to speed up the diffusion process.

---

[1] Ibid., p. 157.

[2] Everett M. Rogers and Rabel J. Burdge, Muck Vegetable Growers: Diffusion of Innovations among Specialized Farmers, Research Circular 94 (Wooster: Ohio Agricultural Experiment Station, 1961).

2. Although muck farmers were small in number in Ohio, their relative importance was suggested by the value of crops grown and by the number of workers employed.

3. It was important to study highly specialized farmers since there is a trend toward increasing specialization of American agriculture as a result of new farm technology. Past research by rural sociologists had suggested that the communication behavior of specialized farmers was much different from that of relatively less specialized farmers.

Data were obtained on forty-four of the forty-nine farms by personal interviews with the 61 operators involved. The fact that several of the farms operated under partnerships, father-son arrangements, and owner-manager combinations accounted for the 61 operators on forty-four farms. Some of the findings related to the forty-four farms; other findings pertained to the 61 farmers. To provide a contrast between specialized farmers and more general farmers, certain comparisons were made against a state-wide sample of 104 commercial farmers who had been interviewed in 1957.

Findings were reported in terms of frequencies and percentage distributions. Eight figures which presented the data pictorially, in bar graphs, or in frequencies and percentages, were used. Comparisons of muck farmers with commercial farmers were reported for the following areas: (1) personal characteristics of growers including age, education, farm size, and cosmopoliteness, (2) communication behavior including growers' sources of information, contact with scientist, and communication with their neighbors about muck vegetable innovations, (3) perception of value of sub-station farm, and (4) perception of future of farming.

It was reported that one of the useful dimensions of the analysis made was that of adopter categories. Farmers were classified into five adopter categories (innovator, early adopter, early majority, late majority, and laggard) on the basis of their scores on an adoption-of-muck-innovation scale. This scale consisted of fourteen recent muck vegetable practices which had been recommended by scientists at the Ohio Agricultural Experiment Station. An adoption score for each farmer was so computed that more points were given for earlier than for later adoption of each recommended practice. The investigators used a method previously used by Rogers to classify forty-four muck farmers by adopter categories.[1] To validate the adopter categories, a professional who had worked closely with the forty-four muck farms was asked to rate subjectively each farm on a five-point scale indicating a tendency to innovate muck farming practices. Intraclass correlation between adoption scores and the judge's ratings was +.44 and Robinson's A was .61. Both were significant at the .01 level of confidence.

---

[1] Everett M. Rogers, "Categorizing the Adopters of Agricultural Practices," Rural Sociology, Vol. 23 (1958), pp. 345-354; and Characteristics of Agricultural Innovators and Other Adopter Categories (Wooster: Ohio Agricultural Experiment Station, in press), cited by Rogers and Burdge, ibid., p. 17.

An analysis was made of the differences among the muck vegetable farmers in terms of adopter categories. These differences related to the following factors:

1. Characteristics--age, education, size of farm, number of family workers per farm, number of hired workers per farm, travel outside of county to observe farm innovations, travel to Florida, and distance between residence and Sub-station Experiment Farm;

2. Communication Behavior--direct contact with Experiment Station as most important source of information, visits to Sub-station in past year, attendance at Muck Crop School, and whether or not named as source of information by four or more other farmers.

Opinion leaders, those farmers named by four or more other farmers as a source of information or advice, were compared to other farmers in terms of the items listed in (1) and (2) above.

The findings may be summarized briefly as follows:

1. Muck farmers were highly research oriented; they made frequent trips to Experiment Stations, to neighboring farms and to out-of-state farms to observe innovations in muck farming.

2. Generally, muck farmers were younger, had more education, employed more labor, and were more cosmopolite in travel than were the average commercial farmers in Ohio.

3. Innovators had more education and larger farms. They did more cosmopolite travel and were geographically closer to the Sub-station Farm than were the laggards. They also had more direct contact with scientists than did the laggards, and were more likely to be named as opinion leaders.

4. The effectiveness of the Sub-station Farm was indicated by the number of farmers who had information-contact with this facility.

5. Muck vegetable farmers generally were more optimistic about the future of the farmer than were the average commercial farmers in Ohio.

It was concluded that muck vegetable growers were more highly specialized, better educated, younger, more cosmopolite in information-seeking behavior, more research oriented, and engaged in more intensive farming than did average farmers. They were also in much closer communication with agricultural scientists than were the average commercial farmers in Ohio.

The research methodology represented in the foregoing two studies has been used extensively by sociologists in conducting adoption-diffusion research. Data were obtained through interviews with respondents. No information was given on the pretesting of the instruments used to collect data. Moreover, for the basic data to be valid, one must assume that each farm operator was able to recall precisely and accurately his experience with each stated practice. However, one may assume that the researchers were skillful in eliciting recall of experiences since they are reputable sociologists.

The data for the Rogers-Burdge study were obtained on forty-four of the forty-nine farms in the population. However, it must be recognized that some bias may have been introduced. This resulted from the fact that the five farms that were not included were not omitted by chance but because of the operation of unknown factors which caused them not to be available. The fact that the operators of these five farms could not be contacted on any of three visits or by mailed questionnaire suggests that they may have differed significantly from those who were included. The extent to which the omission of the five farms biased the forty-four as a sample of Ohio muck growers is not known. Copp, Sill, and Brown reported that theirs was a random sample of 175 dairy farmers in a western Pennsylvania County, but they gave no indication of the percentage of the population which the sample represented. A random sample of this size reduces the sampling error to the point at which sensitive measures of differences may be obtained.

The statistical analyses employed seem appropriate for the purposes used. However, they were not described in detail. Copp, Sill, and Brown reported the confidence levels of the difference between progress scores classified by presence or absence of peer informational sources for the awareness and interest stages for the three practices studied. This information was included in one of the four tables in which the data were presented. Rogers and Burdge reported results of only one statistical analysis in a footnote.[1]

The two studies were reported similarly. For each study the investigators reported an analysis made of the data and the findings, after which they reported other analyses and findings sequentially. In terms of analyzing data and reporting findings, this procedure is acceptable. However, the reader is not adequately informed as to the purposes of the research nor of the analyses to be made of data. For the astute consumer of research such reporting of research leads one to believe that the researchers had assembled a mass of data and that sequent to one analysis, they decided to make another analysis because it would be of interest to know the findings. The basic principles of reporting a piece of research involve clear statements of the purpose(s) of the research and the manner in which data are to be obtained and analyzed. It is evident to the present investigator that such research reports are addressed only to Cooperative Extension Service Personnel for the primary purpose of reporting findings in order to improve the effectiveness of working with constituents rather than reporting a piece of research, per se. In other words, the reports of research, particularly in this category, are not presented for the type of analysis attempted in this study. This fact makes it most difficult for the present investigator to determine the quality of research.

---

[1] _Ibid_., p. 17.

The major contribution of each of these pieces of research is to the practical situation under study. In the strict sense, the findings apply only to the subjects identified in each study. However, the reports of the general characteristics of the subjects make it possible for other investigators to decide with reasonable accuracy whether or not the subjects utilized may be considered a sample of a larger universe which includes the subjects in which these other investigators are interested. In such cases, the results will be generalizable, but such generalizations were not the primary intent of the investigators.

Other studies listed in the Bibliography which are illustrative of this category include those by Beal and Rogers; Beal, Rogers, and Bohlen; Lionberger; Rogers; Rogers and Beal; and Wilkening.[1]

In the foregoing section the two studies in the sample of twenty-two which were classified in Category II, "Informational Sources--Stated Stages of Adoption" were analyzed and evaluated. Reference was made to eight other studies listed in the Bibliography which are illustrative of this category. The subsequent section deals with the remaining seventeen studies which fell in the category designated as "Methods."

---

[1] George M. Beal and Everett M. Rogers, The Adoption of Two Farm Practices in a Central Iowa Community, Special Report 26 (Ames: Iowa Agricultural Experiment Station, 1960), and "Informational Sources in the Adoption of New Fabrics," Journal of Home Economics, Vol. 49 (1957), pp. 630-634; George M. Beal, Everett M. Rogers, and Joe M. Bohlen, "Validity of the Concept of Stages in the Adoption Process," Rural Sociology, Vol. 22 (1957), pp. 166-168; Herbert F. Lionberger, Legitimation of Decisions to Adopt Farm Practices and Purchase Farm Supplies in Two Farm Communities: Ozark and Prairie, Research Bulletin 826 (Columbia: University of Missouri Agricultural Extension Service, 1963); Everett M. Rogers, "Categorizing the Adopters of Agricultural Practices," Rural Sociology, Vol. 23 (1958), pp. 345-354; Everett M. Rogers and George M. Beal, "The Importance of Personal Influence in the Adoption of Technological Changes," Social Forces, Vol. 36 (1958), pp. 329-335, and Reference Group Influence in the Adoption of Agricultural Technology (Ames: Department of Economics and Sociology, Iowa State University, 1958); and Eugene A. Wilkening, "Roles of Communicating Agents in Technological Change in Agriculture," Social Forces, Vol. 34 (1956), pp. 361-367.

Category III: Methods

Of the twenty-two studies selected as a sample of those conducted under the auspices of the Cooperative Extension Service, seventeen investigations were classified in this category. These seventeen investigations were conceptualized as Methods, using the term in its generic sense. The seventeen studies may be ordered sequentially from those which represent the simplest type to those which represent higher levels of research methodology. Four levels of research methodology were represented. Each level is designated as a sub-category of the total Category III: Methods. Nine studies represent the simplest type of research methodology; i.e., one treatment was applied and the results were assessed in relation to participants as an undifferentiated group. These nine studies are reported in Category III-A. Representing a higher level of research methodology were four investigations in which the results of one treatment were assessed in terms of differentiated groups of participants. These four studies are reported in Category III-B. The third group, Category III-C, includes two studies in which two or more treatments were applied and results assessed in relation to the participants as an undifferentiated group. Finally, two studies represent that level of research methodology in which two or more treatments were applied and results were assessed in relation to differentiated groups of participants. These two studies are reported in Category III-D. In a pattern similar to that followed for studies of Categories I and II, the analysis and evaluation of the studies in Category III are reported.

Category III-A

The nine studies representing those in which one treatment was applied and results were assessed in relation to the participants as a single undifferentiated group included seven survey-type investigations in which one set of data was used to report either the changes, progress, or value judgments resulting from one treatment and two in which a control group was used. To illustrate differences in research methodology, two of the most comprehensive single group studies are reported in more detail than are the other five. Also, this group of nine studies embodies two studies for which the research design provided for the collection of data both before and after treatment and for an experimental and control group. One of the two studies which utilized a control group is reported in more detail than is the other since the design of both is essentially similar.

Norby conducted a study to provide information that would be helpful to Extension workers in planning programs at the county level.[1] The two specific objectives were: (1) to identify principles stating conditions which, when met, should result in effective county Extension program planning, and (2) to apply these conditions as criteria for determining the effectiveness of the procedure carried out by the Waupaca County Extension staff and program planning committee in developing a long time program plan. Eight principles were identified, and twenty-seven conditions embodied within the context were deduced, all of which were supported theoretically by statements of guidelines from the professional literature including generalizations from research.

Data were obtained from observations, tape recordings, and six questionnaires. Each of the eleven program planning meetings held during the six-month period and the follow-up work by the Extension staff during the subsequent five-month period were observed. Transcriptions of the tape recordings of the formal portions of three of the four county committee meetings and of the seven area meetings were analyzed for content of discussions. A comprehensive questionnaire was used during a personal interview with each committee member following the first meeting, and a similar one was used subsequent to the last committee meeting. Briefer questionnaires were administered to committee members at the close of each of the first three county planning meetings, and another was administered to members of the County Extension staff at the close of the project to obtain reactions to their experiences in the planning process. The County Program Plan which was developed provided additional data.

---

[1] Oscar Woodrow Norby, "An Appraisal of Long-Time Cooperative Extension Program Planning in Waupaca County, Wisconsin" (unpublished Ph.D. dissertation, University of Wisconsin, 1961).

From an analysis of these data an over-all evaluation of the extent to which each condition had been met was made. A five-point scale--"extremely well,""quite well," "fairly well," "not very well," and "not at all"--was used, and a scale value for each condition was assigned.

The major section of the report presented a descriptive interpretation of the data, appropriately organized under each of the twenty-seven conditions, after which the scale value was assigned. Of the twenty-seven conditions, four received a scale value of "extremely well"; six "quite well"; twelve "fairly well"; five "not very well"; and no condition was scored as "not at all." Finally, a procedure for long-time program planning was submitted. This descriptive procedure, consisting of twenty-four steps, embodied the twenty-seven conditions.

Lintner conducted a study to determine the effectiveness of the Institutional-On-Farm Training Program in Ohio in the light of its objective after five years of operation.[1] The objective of the program was, "to enable a veteran to become established in farming by the end of his training period to the extent that he will be able to support himself and his family under normal economic conditions without government subsistence."[2] Primary objectives of the study were to determine: (1) the veterans' status in farming, (2) the contributions of the Training Program toward this establishment, (3) their desire for future training, and (4) the implications of the Institutional-On-Farm Training for improvement of present programs in adult education.

A schedule consisting of five questionnaires was constructed and pretested. It was mailed to a randomly selected state-wide sample of 300 of the approximately 6,000 veterans who had either completed or interrupted their training. The 259 usable schedules which were returned represented approximately 4.3 per cent of the population. The author reported that informal investigation indicated that there was not more than 5 per cent discrepancy between veterans in the sample who did and those who did not return the questionnaire, but did not make explicit the basis for his statement. Data were tabulated for three groups, "full-time farming," "part-time farming," "not presently farming," and were reported in percentages. The descriptive analysis of data was reported in six lengthy chapters which included the social and economic status of respondents, indication of establishment in farming, contributions of program to veterans established in farming, veterans' desires for further training, implications for future programs in Institutional-On-Farm Training, and implications for improvement of present programs of Adult Education in Agriculture. Of the many findings reported and of the seventeen conclusions drawn, the following is a brief summary.

---

[1] Julius Harold Lintner, "A Study of the Effectiveness of the Institutional On-Farm Training Program in Ohio with Implications for Future Programs in Adult Education" (unpublished Ph.D. dissertation, Ohio State University, 1952).

[2] Ibid., p. 1.

When considering the single objective measure of establishment in farming, i.e., physically farming on either a full-time or a part-time basis, the objective was achieved with 59.4 per cent of veterans who entered into training. Of these 75.3 per cent were farming full-time and 24.6 per cent part-time. Veterans held high regard for the classes and more than 60 per cent felt the need for further educational programs. Practically all were willing to continue in such programs without subsistence payments. Nevertheless, Lintner concluded that any future program should provide both subsistence and instruction and that adult education programs in Agriculture should provide more service for farmers, promote greater coordination between agencies, and devote more time to consideration of individual rather than to group problems.

The remaining five studies are similar to Lintner's study in several ways. Data were obtained either by a mail questionnaire or by a questionnaire used during a personal interview. All data were reported descriptively either in frequencies or in percentages. No statistical tests were employed. The minor differences for each of the remaining five studies are reported briefly in the subsequent paragraphs.

Clem, in addition to obtaining much data from home demonstration agents through a mail questionnaire, devised a method to evaluate the quality of craft work.[1] She arranged for samples of sixteen types of craft articles, the "best" from each of the ten counties under study, to be submitted for judging. Fifty-five articles were rated for quality of design and for technique by 120 students who were currently enrolled in related art and craft courses at the University of Tennessee. The students were considered competent judges because of their unusual rank on a special standardized test administered previously for that purpose. Using a rating sheet prepared by Clem, the judges evaluated each article on the basis of a numerical score of 0-10 as follows: 9-10, excellent; 6-7-8, good; 3-4-5, fair; and 0-1-2, poor. Detailed results from these ratings were reported descriptively, accompanied by six summary type tables. The final composite evaluation of the craft articles was "good."

With the exception of the attempt to evaluate the quality of the craft work, Clem's study was only a status survey. No data were given to indicate the extent of improvement in craft work nor whether this work was a direct result of extension programs, per se.

---

[1] Margaret Louise Clem, "An Evaluation of the Present Status of the Craft Phase of the Tennessee Agricultural Extension Service in Upper East Tennessee To Be Used as a Basis for Future Program Development" (unpublished M.S. thesis, University of Tennessee, 1961).

Holder used two questionnaires, one for urban home demonstration club members and one for nonmembers, during personal interviews with a randomly selected sample of homemakers representing each group.[1] Only the immediate reaction of the club member was obtained as to whether the benefit of selected phases of the home demonstration program had been "much," "some," or "none." The non-club member gave immediate responses to phases of homemaking in which she was most interested. As indicated by Holder, the information obtained was of much value to her in planning future programs. However, the study was a survey of status rather than an evaluation of a program.

Jennings obtained much personal-type information and determined the extent to which homemakers had adopted twelve recommended homemaking practices.[2] Personal interviews were conducted with a random sample of white homemakers in the Parish by eight women, all of whom were employed by the Cooperative Extension Service. This status survey was of value for planning future programs, but the level of evaluation was simply that of reporting without analyzing results.

---

[1] Nadine D. Holder, "Evaluation of the Home Demonstration Club Program in Meeting the Interests and Needs of Urban Homemakers of Washington County, Oklahoma" (unpublished M.S. thesis, Oklahoma State University, 1959).

[2] Sybil Jennings, "Studying Home Demonstration Work

Matott, concerned with finding ways to improve the farm management phase of the Farm and Home Management Extension Program, conducted a study to evaluate the program for the first three-year period it had been in operation.[1] Using a four-sectioned pretested questionnaire, he interviewed thirty-three of the forty families who had originally enrolled and completed the first three-year cycle of the Program. The report contained a lengthy descriptive account of the conduct of the Program. The resulting changes as revealed by the data obtained from the interviews and from analysis of farm records were reported descriptively

---

with Rural Homemakers in Winn Parish" (unpublished M.S. thesis, Louisiana State University and Agricultural and Mechanical College, 1954).

[1] Howard William Matott, "A Study of the Farm Management Phase of the Farm and Home Management Program in Chenango County" (unpublished M.S. thesis, Cornell University, 1958).

with summary tables interspersed. Thus, the study surveyed changes in program participants and provided information which could be used in improving the program.

Reuter, to determine the effectiveness of the project leader training meeting, analyzed opinions obtained from all county home agents of a state in response to a mail questionnaire.[1] The report consisted of a descriptive analysis of the data in terms of the following: (1) methods used by home agents in organizing, planning, and conducting leader training meetings, (2) follow-up practices and evaluation methods used in connection with leader training meetings, (3) types of assistance received by home agents from extension specialists, (4) types of assistance desired from extension specialists to help with leader training responsibilities, and (3) recommendations for improving leader training meetings. The study was a survey of the status of the leader training system in a given state which also elicited additional opinions from home agents for the improvement of the leader training system.

---

[1] Ava Marie Reuter, "An Evaluation of the Leader Training System of Home Demonstration Work in Wisconsin as Viewed by Home Agents" (unpublished M.S. thesis, University of Wisconsin, 1959).

Other studies listed in the Bibliography which appear to have characteristics similar to the foregoing seven studies include those by Carroll, Cochran, Fessenden and Rohrer, Hendrickson, Holloway, Minter, Neilson and Bittner, Preston, Samendinger, and Winterfeldt.[1]

The research methodology employed for each of the two remaining studies represents a higher level than did the preceding seven studies. Because results of only one treatment was assessed in relation to a single undifferentiated group of participants, these two studies were classified in the first sub-group of studies under Category III: Methods. However, the use of a control group against which treatment effects were compared represents a significant improvement over the simple design which reports progress and changes made by participants only during a given period. A report of the two studies follows.

---

[1] C. B. Carroll, "A Normative Survey of Viewers of the County Agent's Daily Television Program in Caddo Parish, Louisiana" (unpublished M.S. thesis, Louisiana State University, 1959); Lacy E. Cochran, "A Study of the Progress of a Group of Veterans Enrolled in Institutional On-Farm Training in Hardy County, West Virginia" (unpublished M.S. thesis, West Virginia University, 1954); J. G. Fessenden and W. C. Rohrer, A Study of an Urban Home Economics Extension Program, Baltimore, Maryland, Miscellaneous Publication 159 (College Park: University of Maryland Agricultural Extension Service, 1957); V. C. Hendrickson, "An Appraisal of the Rural Resource Development Program in Price County, Wisconsin" (unpublished M.S. thesis, University of Wisconsin, 1960); M. E. Holloway, An Evaluation of the Maryland Extension Fair Exchange TV Program of September 17, 1960 (Baltimore City Extension Service, 1961); E. M. Minter, Money Management Workshops for Young Homemakers--A Study in Three Indiana Counties, Research Bulletin 731 (Lafayette: Indiana Agricultural Experiment Station, 1961); James Nielson and R. F. Bittner, Farm Practice Adoption in Michigan, Technical Bulletin 263 (East Lansing: Michigan Agricultural Experiment Station, 1958); J. C. Preston, "An Evaluation of the Farm and Home Management Program in Genesee County, New York" (unpublished M.S. thesis, University of Wisconsin, 1959); H. Samendinger, "The Cornell Farm Study Course Program" (unpublished M.S. thesis, Cornell University, 1955); E. P. Winterfeldt, "Evaluation of the Home Demonstration and 4-H Girls Programs in Pittsburg County, Oklahoma, Based on Opinions of Home Demonstration Club Officers" (unpublished M.S. thesis, University of Wisconsin, 1959).

Alexander and Longest conducted a comprehensive evaluation study over a four year period of the farm management phase of the Farm and Home Management program in New York State.[1] The major purpose of the study was to compare changes in farm practices and related knowledge, and in income and related business factors, of participants in the program with those of a control group of nonparticipants.

---

[1] Frank D. Alexander and James W. Longest, Evaluation of the Farm Management Phase of the Farm and Home Management Program in New York State, Extension Study No. 1 (Ithaca: New York State University Extension Service, Cornell University, 1962). This report and one of the mimeographed reports in the evaluation series, "Evaluation Study of the Farm and Home Management Program in New York State--Design and Methodology," Report No. 9 (Ithaca: Office of Extension Studies, Roberts Hall, Cornell University, September 30, 1961) were read for this analysis.

The program was initiated in July, 1954, and by the spring of 1956 when the evaluation study was launched, twenty counties had inaugurated farm management programs. This intensive educational program focused on management of the farm business. The program implemented this focus by teaching the techniques of keeping and of systematically analyzing farm business records followed by individual and group counseling on ways to correct such weaknesses as were revealed in the analysis of records.

It was specified early in the report that a longitudinal study of this type involves technical difficulties which can prevent final completion. The difficulties encountered involved the sample, the criterion measurement, and two natural variables, the control of which was either not preferred or not possible. In this brief analysis each of these difficulties and the decisions made as how best to deal with them are reported.

Data used to compare changes of participants in the program with those of a control group of nonparticipants were obtained during the 1956 initial and the 1959-1960 terminal surveys of participants and of nonparticipants. Trained interviewers used two survey instruments in the conduct of personal interviews. The instruments were: (1) a specially constructed and pretested questionnaire designed to secure information of farm practices, general family characteristics, and personal characteristics of the farm operator, and (2) Cornell Labor Income Blank #40, an instrument developed by the Cornell Department of Agricultural Economics for recording farm business data.

The original sample was composed of 250 participants and 150 nonparticipants who were selected as follows. First, on the basis of six criteria, ten of the twenty counties were selected for inclusion in the study. Twenty-five program participants were randomly selected from each of six selected counties, and all participants (25) in each of the four other selected counties were included. To provide a control group, the 25 participants in six of the ten counties were paired with 25 nonparticipants. Matching was done by county agricultural agents on the basis of nine variables. Thus, in the 1956 initial survey there were 150 matched pairs of participants and nonparticipants in six counties plus an additional 100 participants in four other counties. When tested by chi square, there was no significant difference between the sample of participant operators and the finite universe of participants on seven selected variables. Therefore, the sample of 250 participants was considered to be reasonably representative of the universe of 709 participants.

During the period of the study, the sample of 250 participants experienced a total attrition of 18 per cent, a loss of 46 members, and the control group of 150 nonparticipants had a total loss of 28 per cent, 43 members. The primary reason for the higher attrition for the control group was that 11 per cent (17) of the 150 nonparticipants became participants in the program in the interim between the initial and final surveys. According to the investigators this fact illustrated the importance of having experimental control over the control group if such could be achieved, and furthermore, it was possible that this type of attrition biased the remaining group more than did other forms of attrition. Consequently, the measurement of the changes in the stated phenomenon was based upon the 204 participants and the 107 nonparticipants who were available at the conclusion of the study. Because the attrition in both groups was of the magnitude to prevent utilization of the original matching, it was necessary to attempt rematching of the remaining 107 nonparticipants and 204 participants. Since ample quantitative data were available from the initial survey, pair-matching was chosen and eleven variables were used in the trial matching. When the final trial matching, involving 87 participants and 87 nonparticipants, was subjected to a t-test for nonsignificance, the matching was found to be valid for eight of the eleven variables. Therefore, these 87 matched operators were used for testing the differences in changes made by the participants compared to those made by the nonparticipants.

The relationship between the 87 participants and the remaining 163 of the original sample of 250 was tested by means of the t-test for each of the eleven matching variables to determine whether or not the two sub-groups differed significantly. It was found that they did differ significantly on six of the eleven variables. Therefore, the 87 matched participant operators could not be considered to represent the original sample of 250 participants; hence, they did not represent the universe of participants. It was recognized that conclusions resulting from a comparison of the 87 matched participants and nonparticipants would not represent the total effects of the program. But, since the total sample of participants could not be matched, such conclusions would be the best estimate of the effectiveness of the program. According to the investigators, it was "far better to have the increased control attained by use of a pair-matched participant and control group, whether the matched participants do or do not represent all participants in the original sample, than to use all participants and compare them to a control group with which they would be poorly matched."[1]

After the initial survey was made, seven farm knowledge and practice tests were developed from items listed in the questionnaire and from the labor income record forms. These tests related specifically to five subject-matter fields. The tests were: (1) farm management knowledge and practices, (2) dairy feeding practices, (3) dairy breeding practices, (4) dairy disease control practices, (5) agronomy practices for corn, (6) agronomy practices for oats, and (7) agronomy practices for hay and pasture. Knowledge and practice scores were calculated for each of these tests, and were used as measures of the extent to which each respondent was following recommended practices in the respective subject-matter areas. The procedures followed to develop these tests and to determine validity, reliability, and discriminatory power of items are reported briefly in the subsequent paragraphs.

---

[1] Alexander and Longest, "Evaluation Study of the Farm and Home Management Program in New York State--Design and Methodology," op. cit., p. 40.

One or more agricultural specialists in each of the five subject-matter fields had assisted in the preparation of the questionnaire. The same specialists were asked to:

1. Indicate the practices that would apply to all dairy farm operators;
2. Indicate those questions for which the answers would differentiate adequately among operators on recommended management knowledge or behavior;
3. Weight the retained questions so that those most crucial to good management received the highest number of score points; and
4. Assign appropriate values to each of the multiple answers for those questions with multiple responses.

In order to compare different score levels achieved on different practices by the same and different operators, the raw scores received by respondents on each of the seven tests were converted to per cent of total possible score. After the score values were assigned, scores for a number of farm operators from each of the ten counties were calculated. These specialists plus one or more specialists from each of the subject-matter fields ranked about ten of the same operators on management in their respective subject-matter fields. The rankings of these specialists were compared to one another, and to the rank of the calculated score for the operator. When differences of any significance occurred among the rankings by the specialists or between calculated scores and ranks by the specialists, the ranks and score weights were re-examined. Specialists were able to reach unanimous agreement as to acceptable weights.

Because of the lack of an outside criterion with which to correlate the test scores, the validity of these tests could not be ascertained in that manner. However, face validity was assumed since the cooperating subject-matter specialists were competent to specify the content of their fields relevant to management of a dairy farm. Too, there was general agreement by specialists in the respective fields other than those who prepared the original questions that the items did test relevant subject matter.

Reliability was determined by the internal consistency method. It was acknowledged that the retest method would have been the most appropriate method, but neither this method nor the alternate forms method was possible because the test items for scoring were not selected from the data collecting instruments until after their initial administration. Besides the cost was beyond limits of the study budget. The Kuder-Richardson formula #20 was used for all tests except the three agronomy tests for which Dressel modification of Kuder-Richardson 20 was used. The Spearman-Brown Prophecy formula for predicting reliability of tests for assumed specified lengths was applied to the reliability coefficients to determine reliability coefficients where the assumed length was set as twenty-five items for all seven tests. To apply these formulas it was necessary to rescore the test items on the basis of a dichotomous pass-fail split. For test purposes, therefore, all items in all the tests, except the agronomy tests, were weighted equally.

The reliability coefficients obtained were low for three tests, the dairy disease control, oat, and corn practices tests; the four other tests had only moderately low reliability coefficients. The results from the Spearman-Brown Prophecy formula indicated that if additional items had been added to the shorter tests, the coefficients would have improved considerably. Acknowledging that the low reliabilities of the tests were crucial, particularly when they were to be used to measure differences, the investigators concluded that the use of the test scores as measures of the operators' practice levels in general subject-matter areas were preferred to the use of numerous separate knowledge and practice items. They pointed out three factors which had contributed to the low reliability coefficients:

(1) relatively few items in the tests, (2) a relatively homogeneous group tested, and (3) heterogeneity of factors measured in any one test.

In addition to the judgments of the specialists concerning discriminatory power of items, the D-index test was applied to the initial survey data for 400 respondents, consisting of 250 participants and 150 nonparticipants. Items were eliminated from tests on the basis of the D-indexes and the percentage of operators who passed the items. However, some items which were low on discriminatory power according to the D-index and were high on percentage of operators who passed them were retained. This was done only if the item were one which was thought to be discriminatory when applied to a random sample of all commercial farmers in the state rather than to this group of operators who were in general above the average of the operators in the state on several important variables.

The report contained an elucidative discussion of these procedures, reinforced by statistical formulas and by source references used, and by numerous tables presenting all results of statistical analyses.[1]

---

[1] Ibid., pp. 9-18; 41-49.

The authors pointed out other difficulties encountered which the present investigator designated as natural variables. Program agents did not work with every participant to improve all specific management practices. The range was from almost all participants who were exposed to some of the recommended practices to a few or none who were exposed to others. Except for the general focus in all counties upon record keeping, business analysis through the use of these records, and a few other practices, the agents involved in the program differed with respect to the subject matter emphasized. Therefore, the differences in change from initial to terminal scores for the experimental and control groups could not be expected to be as great as it might have been had the program and the tests of its results focused on a more limited number of practices. However, this did not imply that the program should have had a narrower focus; rather, to the extent that analysis of records revealed weaknesses with which the agents assisted the participants who had these weaknesses, the over-all influence of the program may have been greater than if agents arbitrarily had taught all participants selected knowledge and practices regardless of their individual needs. One other factor might have effected the extent of change. The members of the experimental and control groups were from the same counties, and in these counties general information in relation to the farm businesses of the participants was made available by the Cooperative Extension Service through various mass media. Therefore, it was most unlikely that the nonparticipants completely escaped from some program influences.

All data were analyzed statistically. The difference in change for each pair of operators was determined as follows. First, the difference between the 1956 and 1959-1960 values for each participant and for each nonparticipant was calculated. Secondly, the difference between the change of each participant and his matched nonparticipant was found. Then, the mean difference in the change of participant and nonparticipant pairs was tested for significance by using the one-tail t-test which was considered more discriminating than a two-tail t-test for testing change. The .05 level of confidence or above was accepted for determining statistical significance.

An analysis of the comparative changes was presented in considerable detail. For the most part such changes favored the participants. Only the following were statistically significant:

1. The gain between 1956 and 1960 in the percentage scores on the farm management knowledge and practices test for the eighty-seven participants was significantly greater than for the eighty-seven nonparticipants.

2. The eighty-seven participants showed a significant gain over the eighty-seven nonparticipants on five of the twenty-five internal items of this test. On one of these internal items the participants and nonparticipants were not matched in 1956, and the nonparticipants had a significantly higher score on this item than did the participants.

One conclusion was drawn: The program had a favorable influence on the participants.

Holik and Moore conducted a longitudinal study which was similar to the Alexander-Longest study in one major aspect.[1] Data were obtained from participants and from nonparticipants during interviews conducted approximately eight years apart, 1948 and 1956. It differed in several aspects. First, it involved the program in only one county. In 1948, fifty farmers were randomly selected from the seventy-eight who had participated in the program. For each of these farmers two neighboring nonparticipating farmers, with about equal cropland acreage, were selected. Thus, farmers of Group II, the experimental group, were matched approximately with farmers of Group I, the control group, on this variable; no assumptions were made that they were comparable in characteristics other than that of farm size. The report gave no specific information on the development or pretesting of the questionnaires, except to state that the two which were used were similar and that some additional information was obtained in the second survey. Data were analyzed to compare the two groups of farmers and the report contained a lengthy description of these comparisons. Thirty-eight tables, presenting data in frequencies, percentages, and means, were interspersed within the context. It was not clear as to the extent to which differences had been tested for statistical significance. Frequently, in the context such statements as, "This difference was found to be statistically significant," or "These differences were not statistically significant," were made. At one point reference was made to a chi-square test, the results of which were significant at the .05 level. Neither a summary of findings nor of conclusions was presented; rather, findings and conclusions were interspersed within the whole report. Essentially, the conclusion was summarized in the statement, "The data suggested that those farmers who continued to participate in Balanced Farming profited from the

---

[1] John S. Holik and Clay R. Moore, "Balanced Farming Evaluation Study" (University of Missouri Agricultural Extension Service, 1957).

Program."[1]

It is the judgment of this investigator, based on a comprehensive search for evaluation research in Cooperative Extension, on examination of a large number of summary reports of the research, and on the research literature in general, that the study by Holik and Moore was the first study to evaluate a longitudinal Program in which an effort was made to compare changes of participants of an experimental group with those of a control group. This pioneering effort served as a basic framework, on which others have improved, and it evidently has influenced the introduction of experimental designs for evaluating long-term Cooperative Extension Programs.

Other studies listed in the Bibliography which are similar to the two previous studies include those by Johnson and Wilkening, Marsh, and Nielson and Crosswhite.[2]

---

[1] *Ibid*., p. 28.

[2] Donald E. Johnson and Eugene A. Wilkening, *Five Years of Farm and Home Development in Wisconsin: A Comparison of Participating and Control Families*, Research Bulletin 228 (Madison: University of Wisconsin Agricultural Experiment Station, 1961); C. Paul Marsh, *An Evaluation of the Farm and Home Development Approach to Agricultural Extension Work in North Carolina*, Extension Evaluation Studies No. III (Raleigh: North Carolina State College Agricultural Extension Service, 1962); and James Nielson and William Crosswhite, *The Michigan Township Extension Experiment*:

Category III-B

This category includes the four studies in which the results of one treatment were assessed in relation to differentiated groups of participants, thus differing from the preceding sub-category in which results were assessed in relation to undifferentiated groups of participants. Since these four studies vary in terms of research sophistication, two of the studies are reported in detail to illustrate the differences in the quality of the research. The similarities and differences of the two remaining studies are noted.

Wilson and Moe conducted one of the earlier studies to evaluate a series of method demonstration television programs, entitled "Let's Make a Dress."[1] A personal interview questionnaire, consisting of fifty-seven items, was constructed "to meet the objectives and the detailed specification for the study."[2] The questionnaire was pretested among members of the same population as that from which the sample

---

Changes in Agricultural Production, Efficiency and Earnings, Technical Bulletin 274 (East Lansing: Michigan State University Agricultural Experiment Station, 1959).

[1] Meredith C. Wilson and Edward O. Moe, Effectiveness of Television in Teaching Sewing Practices, Federal Extension Service Circular 466 (Washington, D. C., 1951).

[2] Ibid., p. 3.

was drawn. A random sample of 251 homemakers was drawn from the 974 who, before the program opened, had requested by telephone a copy of the program bulletin, "Making a Dress at Home." Using this questionnaire, interviews were conducted with this sample by experienced interviewers within three to four weeks after completion of program.

The investigators found that of those who had viewed even one program nine out of ten women said they had learned new things. Almost one-half of the women used ideas learned, and the proportion of women using this practice increased with the number of programs viewed. Ten summary statements and conclusions were drawn of which the following are representative:

> 1. An interesting characteristic of the audience studied was the high proportion of young mothers and mothers of young children, a segment of the population that has been difficult to reach with the usual methods employed in adult education. . . .
> 2. The data point emphatically to the importance of a bulletin or similar printed matter as a supplement to television demonstrations in teaching dress making practices. . . .
> 3. These findings indicate the effectiveness of TV as an educational medium. . . . A similar series could be readily prepared by the extension staffs of the State Colleges.[1]

---

[1] *Ibid.*, pp. 1-2.

Guarino conducted a study which is similar in terms of the means of obtaining and reporting data.[1] In this study, as in the previous one, an interview schedule was developed and Guarino pretested the schedule with five dairy operators. The subjects used were not randomly selected; instead, they were simply 118 of the 124 qualified subjects who constituted the county population. The subjects were classified into four attendance categories. Thirty-four were classified as "regulars" (attended meetings six to twelve times a year), sixteen as "occasional" (attended meetings three to five times a year), twenty as "seldom" (attended meeting once or twice a year), and forty-eight as "never." The percentage of each attendance category who (1) used selected extension education media, (2) adopted recommended practices, and (3) possessed specified socioeconomic characteristics was reported.

Other studies listed in the Bibliography which are illustrative of the type of research methodology employed in the two foregoing studies include those by Eachler, Dell, and Alexander; Erickson and Olien, La Frenier, Lionberger, and Welker.[1] According to the summary reports for one of these studies, two sets of data were obtained, before and after treatment, and the differences in knowledge scores were analyzed statistically.[2]

---

[1] Russell Guarino, "Evaluation of Monthly Dairy Discussion Group Meetings as an Agricultural Extension Teaching Method in LaFayette Parish" (unpublished M.S. thesis, Louisiana State University and Agricultural and Mechanical College, 1961).

The remaining two studies in Category III-B utilized a control group. One is reported in detail and the similarities and differences between the other study and the reported study are noted.

Rogers and Havens tested the effect of result demonstrations in changing attitudes about the use of fertilizer.[3]

---

[1] R. E. Eacher, J. C. Dell, Jr., and F. D. Alexander, *Evaluation Study of the TV Dairy Cattle Feeding School*, Special Report 6 (Ithaca: Cornell University Agricultural Extension Service, 1962); S. Erickson and C. Olien, *Evaluation of TV Clothing Series "What Shall I Wear?"* (St. Paul: Minnesota University Agricultural Extension Service, 1961); B. M. La Frenier, "A Study of the Effectiveness of Television in Teaching Clothing Construction" (unpublished M.S. thesis, University of Tennessee, 1960); Herbert F. Lionberger, *Television Viewing in Rural Boone County*, Bulletin 702 (Columbia: University of Missouri Agricultural Experiment Station, 1958); Walter Welker, "The Cooperative Extension Television Audience in a Pennsylvania Rural Area" (unpublished M.S. thesis, Pennsylvania State University, 1962).

[2] Eachler, Dell, and Alexander, *op. cit*.

[3] Everett M. Rogers and A. Eugene Havens, *The Impact of Demonstrations on Farmers' Attitudes toward Fertilizer*, Research Bulletin 896 (Wooster: Ohio Agricultural Experiment Station, 1961).

The objectives were to determine:

1. The extent of change in attitudes toward fertilizer as a result of a demonstration program;
2. The extent of changes in knowledge about fertilizer and use of fertilizer as a result of a demonstration program;
3. The extent to which attitudes toward fertilizer affect the relationship between knowledge and use of fertilizer; and
4. The characteristics of those farmers with attitude change.

The research design provided for a before-after study of subjects in an experimental group and utilized a control group. On the basis of similarity two adjacent counties were chosen as the study areas. An intensive fertilizer demonstration program was conducted in the treatment county and no demonstrations were held in the control county.

Random area samples in both the treatment and control counties were selected and all farm homes in the designated areas were visited in December, 1958. All farmers operating twenty acres or more who could be contacted in three visits were interviewed. In the treatment county interviews were completed with eighty-six farmers. According to the 1955 census this number represented 3.8 per cent of all farms in that county. In the control county forty farmers were interviewed. This was 2.1 per cent of all farms. Again, during March and April of 1960, interviews were conducted with seventy-seven of the original eighty-six farmers in the treatment county, and with thirty-seven of the original forty farmers in the control county.

During the 1959 crop year, forty-seven farmers in the treatment county carried on meadow and corn fertilizer demonstration plots. In the development of these plots farmers adhered to several recommended practices, the most important of which pertained to the amount of fertilizer used. For the corn plots fertilizer was applied at three levels: (1) the farmers' normal rate of fertilization, (2) the amount recommended as a result of soil tests, and (3) the soil test recommendation levels plus fifty pounds of nitrogen per acre. For the meadow plots fertilizer was applied at two levels: (1) the farmers' normal rate of fertilizer, and (2) the amount recommended as the result of soil tests. Related educational work carried on during the year consisted of posting the plots, conducting tours to selected plots at various times during the year, and presenting the results of the demonstration plots at a banquet. The demonstrator farmers also were interviewed after completing their demonstration plots. These interviews were conducted during December, 1959, before most meetings had been held to discuss the results.

During the interviews data were obtained on existing attitudes, knowledge, and use of fertilizer, attitudes toward demonstrations, and other characteristics of the farmer. An objective measuring device, an "attitude toward fertilizer scale" consisting of eight items, was administered during both interviews. Five items of the scale had been used previously by other researchers. Six of the eight items were of a sentence completion nature. The investigators stated:

> This type of projective technique was utilized in order to secure a more accurate response. Projective items are generally regarded as one method of avoiding bias from socially acceptable answers.[1]

Reliability of the scale was determined by the test-retest method (r = +.49). Unidimensionality, determined by the Guttman technique was 78 per cent which was less than the 90 per cent usually required. Internal consistency ranged from +.25 to .66, as measured by item-to-total-score correlations. Knowledge of fertilizer was measured by a six item scale. The report gave no further information on this scale other than to cite two typical items: "If the price of a mixture of 3-12-12 is $50 per ton, and the price of 6-24-24 is $80 per ton, which is the best buy?" and "How may potash deficiency symptoms be recognized?"[1]

---

[1] Ibid., p. 15.

Three judges scored the item responses from the attitude scale as either favorable (three points), neutral (two points), or unfavorable (one point). Scales from both administrations were scored at the same time with identical criteria. Robinson's measure of agreement ("A") was used to determine the agreement among judges. The results were generally high (.85, .91, and .94).

All data were analyzed statistically and reported in ten tables. The .01 and .05 confidence levels for significance were used. The major analyses and resulting findings are summarized in the following statements:

1. When the average scores obtained from the attitude toward fertilizer scales were tested by a two-way analysis of variance, no significant differences were found between farmers' attitudes before the treatment and after the treatment, or for interaction between counties and years.

2. Knowledge about fertilizer increased in both counties, but there was no significant difference between counties.

3. Twenty-five farmers in the treatment county had a positive attitude change ranging from one to five points higher in the attitude toward fertilizer score. Significant positive relationships between attitude scores and both education level and

---

[1] _Ibid_., p. 22.

classification as early adopters were found. Thus, those farmers whose attitudes became more positive had more education and tended to adopt farm innovations earlier.

4. Significant relationships were found between the three variables, attitudes, knowledge, and use. When partial correlation was computed to determine if knowledge affects the relationship between attitudes and use of fertilizer, the relationship between attitudes and use decreased from +.49 to .27. Since the difference in the two correlations was statistically significant at the .01 level, it was concluded that knowledge about fertilizer significantly affected the relationship between attitudes and use.

5. Soil testing was tested as an intervening variable between attitudes and use of fertilizer. No relationship was found between attitudes and use of fertilizer for farmers who soil test; however, for those who do not soil test, a significant relationship was found between attitudes and use.

The investigator concluded that this study had shown that attitudes, use, and knowledge of fertilizer are highly interrelated, and that the study suggested a dynamic factor, knowledge about fertilizer, in changing fertilizer usage.

The basic research design of the study conducted by Lacy was similar to the previous study.[1] It differed in three aspects, the sample, the precision of developing instruments, and analysis of data. Since Lacy was concerned with changes in attitudes and knowledge as a result of participation in extension program planning, he used the twenty-eight members of the program planning committee as his subjects. This was the experimental group. The twenty-eight members of the control group were arbitrarily selected from the same respective geographical areas as were those of the participating group. The groups were matched approximately on four variables.

---

[1] M. P. Lacy, "The Effects of Involvement on the Participants in Cooperative Extension Program Planning in Waupaca County, Wisconsin" (unpublished Ph.D. dissertation, University of Wisconsin, 1961).

He used a Likert-type attitude scale to ascertain the attitude of subjects. Scores on the extent to which respondents agreed or disagreed with nine statements were computed. Scores on the knowledge test represented the number of correct responses out of a possible sixteen. The complete interview schedule in which these instruments appeared had been pretested, but the report contained no information of further testing of instruments. Moreover, he obtained much other data from interviews, observations, and tape recordings which were used to explore the relationship between members' participation in program planning and seven variables, two of which were attitude toward the Cooperative Extension Service and knowledge of its services.

Lacy chose not to make tests of statistical significance giving as his reasons the inadequacy of his sampling methods and the characteristics of his data. His interpretations were based on comparisons among several distributions of data, percentages, rank orders, and means. Qualitative responses were arbitrarily assigned weighted scores and these scores were the basis for comparative statements interpreting the findings in certain areas.

The investigator was unable to identify other studies listed in the Bibliography which are illustrative of the studies by Rogers and Havens and by Lacy.

Category III-C

This category includes the two studies in which the results of two or more treatments were assessed in relation to the participants as an undifferentiated group, thus differing from the two preceding sub-categories in which one treatment was applied. One of the two studies is reported in more detail than is the other since the design of both is essentially similar. Similarities and differences of the second study from the reported study are noted.

Marsh and Hyman compared participation in programmed self-instruction and in a workshop as methods of increasing

Extension Agents' knowledge of radio in Extension teaching and in changing their attitudes toward its use in Extension teaching.[1] The specific objective was to test the following hypotheses:

1. Programmed self-instruction is as effective as the workshop in increasing the agents knowledge of radio as an Extension teaching method. This is true regardless of length of tenure of the agents.
2. Programmed self-instruction is as effective as the workshop method in developing more positive attitudes toward the use of radio. This holds regardless of length of tenure.
3. Programmed self-instruction changes agents' self-concept of their ability to do radio teaching as much as does the workshop method. This is true regardless of length of tenure.
4. Agents participating in programmed self-instruction in radio will feel that such self-instruction should be widely used in inservice training in other subject matter areas. However, there will be an inverse relationship between favorableness toward the use of programmed self-instruction and length of tenure of the agents.[2]

The basic research design was that of comparing the changes which took place in those agents who participated in self-instruction with the changes in agents who received instruction in a workshop.

---

[1] C. Paul Marsh and Theodore H. Hyman, Programmed Learning in Extension Training: A Comparison of the Effectiveness of Programmed Self-Instruction and Workshop Instruction, Technical Bulletin 161 (Raleigh: North Carolina Agricultural Experiment Station, 1964).

[2] Ibid., pp. 4-5.

Participants for each of the sample groups were selected as follows. From a reliable list of extension agents who were involved in radio teaching two or more times monthly, lists of "new" agents and of "experienced" agents were prepared. New agents were defined as those employed between June 30, 1961, and December 30, 1962; experienced agents as those employed before January 1, 1958. Each of these lists was sampled (with stratification by sex) by the use of random numbers. The researchers explained that because of the higher rate of turnover in home economics personnel, it was necessary to stratify on the basis of sex so that the new sample would not be made up predominately of women and the experienced sample predominately of men. Thus, of each sample of thirty, fourteen were home economics agents and sixteen were agricultural agents. After the total sample had been selected, one-half of each tenure by sex group were randomly assigned to the self-instruction group and one-half to the workshop group.

The programmed material, the knowledge tests, and the attitude scales developed by Tait and Hatch were used.[1]

---

[1] Elton B. Tait and J. Cordell Hatch, "Make Radio Work for You" (University Park: Pennsylvania State University Agricultural Extension Service), cited by Marsh and Hyman, ibid., p. 5.

The programmed material was used by those in the self-instructional group, and the same material was presented to members of the workshop group. Identical pre- and posttests were administered to each group.

Each of the agents assigned to the self-instruction group met the investigator and one of the radio specialists in small groups of from six to eight in a location convenient to his county during a given week. Pretests were administered and the agents received instructions in the use of the programmed material. They completed the program while on the job during the subsequent two-week period after which they returned to the same location for posttests. At this time an attitude scale on the use of programmed self-instruction in Extension training also was administered.

Agents assigned to the workshop group assembled in the state capitol for the two-day workshop training. Actual instruction time was 9.5 hours for the workshop. The average time required for the self-instruction group to complete the program was 6.18 hours.

The data assembled from test scores were analyzed statistically. Either the t-test or the F-test was used to test significance of differences, and the .05 level of confidence was used in designating same. It was reported that where the differences between pre- and post-test scores were tested, the individual pretest and posttest scores were paired in order to eliminate the correlation effect.

A presentation of the analyses and findings constituted the longest section of the report. Seventeen tables offered the supporting data to which frequent references were made. Briefly, the major findings included the following:

1. There were no statistically significant differences between the self-instruction group and the workshop group or between the new agents and the experienced agents in their pretest scores. Thus, each group was at approximately the same level before training.
2. After training, both groups made statistically significant gains on each of the three tests, knowledge, attitude, and in the agents' confidence in their ability to teach by radio.
3. There was no statistically significant difference in gains made by those in the self-instructional group and those in the workshop group on any of the three tests administered.
4. The new agents made significantly greater gains in knowledge than did the experienced agents, and this difference was greater among those participating in self-instruction.
5. On the "attitude toward self-instruction scale," no statistically significant difference was found between the scores of the new agents and the experienced agents.

The investigators concluded that hypotheses one, two, and three were supported by the data. "That is," they stated, "programmed self-instruction proved to be as effective as workshop instruction in increasing knowledge of radio, in developing more positive attitudes toward the use of radio in Extension, and in changing the agents' concept of their ability to do radio teaching."[1] However, the hypothesis that new agents would be more favorable to self-instruction than were experienced agents was not supported by the data. Based on the time required for instruction alone, the investigators reported "that self-instruction was substantially more efficient than was the workshop," and "if time required for travel and the cost of travel and subsistence were also considered, the advantages of self-instruction would be far greater."[2]

Robert attempted to evaluate three teaching methods (the neighborhood meeting, the farm visit, and the circular letter) in getting a recommended farm practice adopted in a given county.[3] The conduct of the study differed from the Marsh-Hyman study in several aspects. First, the selection of the three samples of rice farmers was not controlled to the extent that they could be considered representative of the population. Sampling errors were acknowledged and data were obtained on socioeconomic characteristics and on production practices in order to describe each group more completely. Only one set of data was used. These data were secured by personal interview, following the treatments. Findings were presented in percentages and were reported descriptively. Robert concluded that the data indicated that the two personal contact types of methods, the neighborhood meetings and the farm visits, were more effective in influencing the adoption of the recommended practice than was the circular letter.

---

[1] Ibid., p. 14.

[2] Ibid., p. 16.

[3] James Joseph Robert, "An Evaluation of Selected Teaching Methods in Getting a New Farm Practice Adopted in St. Landry Parish, 1960" (unpublished M.S. thesis, Louisiana State University, 1961).

Other studies listed in the Bibliography that fall in this sub-category include those by Coolican, Donahoo, Schaeffer, Schwartz, Tait, and Willsey.[1]

---

[1] P. M. Coolican, "A Study of the Effectiveness of Teaching by Television Versus Teaching by the Use of an Extension Bulletin" (unpublished M.S. thesis, Michigan State University, 1960); Alvin William Donahoo, "Differential Attitudes of Farmers Toward the Grain Marketing System" (unpublished Ph.D. dissertation, University of Minnesota, 1954); A. J. Schaeffer, "A Study of the Comparative Effectiveness of Three Communicational Channels Used by a Cooperative Extension Agent in Teaching Homemakers" (unpublished M.S. thesis, Michigan State University, 1960); L. M. Schwartz, "A Comparison of Traditional and Mechanical Aid

Category III-D

This category includes the two studies in which two or more treatments were applied and results were assessed in relation to differentiated groups of participants, thus differing from the preceding sub-category in which results from two or more treatments were assessed in relation to undifferentiated groups of participants. Because of the differences in the basic content and the fact that statistical analyses vary in terms of sophistication, both studies are reported.

Porter and Wilson reported a study evaluating the St. Joseph Consumer Marketing Program.[1] The primary purpose of the study was to determine the effectiveness of a special

---

Training Methods in an Extension Home Economics Situation" (unpublished M.S. thesis, Michigan State University, 1961); Eldon B. Tait, "A Comparison of Self-Instructional and a Workshop as Methods of Training New Hired County Extension Personnel in the Subject of Radio Broadcasting" (unpublished M.S. thesis, Cornell University, 1962); and Frank Richard Willsey, Jr., "An Experimental Study of an Adult Learning Situation Involving Three Levels of Training in the Group Discussion Process" (unpublished Ed.D. dissertation, Indiana University, 1962).

[1] Ward S. Porter and Lorene S. Wilson, Evaluation of St. Joseph Consumer Marketing Program, Extension Circular 784 (Columbia: University of Missouri Agricultural Extension Service, 1962).

week-long educational program dealing with pork roast. The message content, consisting of ten specific informational items about pork, was disseminated through the same mass media channels as those regularly used for disseminating consumer marketing information. These channels were radio, television, and newspapers. As a regular part of the program a set of steel meat skewers was offered free to those who wrote in for them.

A cross-section sample of 300 persons was randomly drawn from the St. Joseph telephone directory. Using a uniform schedule developed specially for the study, local volunteers conducted telephone interviews with 246 persons of the sample who did most of the meat planning and preparation in their households. These volunteers had been trained by the professional Extension staff and were familiar with the purpose of the study, the nature of the program, and with the questionnaire. They also had conducted practice interviews. Those who wrote in for the skewers were sampled also and interviewed in the same manner. One week elapsed between the completion of the program and the beginning of the telephone interviews.

The schedule elicited information regarding employment status, age, education, and general knowledge of the

Consumer Marketing Program as presented by each of the three media. Too, it contained five questions which covered the subject matter of the special program and five questions related to pork consumption patterns. Thus, the effectiveness of the special program on pork was measured both in terms of levels of knowledge about the subject content and of estimates of pork consumption patterns. In the absence of any pre-program data, changes in knowledge and in consumer behavior were measured by comparing those of the "exposed" group with those of the "unexposed" group. The exposed group included those who had either seen, heard, or read about the week's educational message on pork.

Findings were reported in frequencies and in percentages. It was reported that no attempt was made to test all differences in percentages for statistical significance, but that such tests had been made whenever there was reasonable doubt as to the level of significance. The .05 level was used.

The appendix to the report contained additional information on the study and sample design, explanations of the statistical procedures employed, and a copy of the telephone schedule. Otherwise, more than three fourths of the report was devoted to the findings which were presented descriptively, supported by nine tables and seven bargraph figures. Organized in sections, the first reported on such characteristics of the sample as education, age, and employment outside the home. The second section dealt with the availability and use of each of the three mass media. The third section reported the exposure to the Consumer Marketing Pork Program and included: (1) number and characteristics of the audience, (2) communication channels, (3) the newspaper audience, (4) the television audience, (5) the radio audience, and (6) a summary of mass media coverage. The final section presented findings on the effectiveness of the special program dealing both with levels of knowledge and pork consumption.

Selected from the findings are the following which seem of most importance:

1. Those who wrote in for the skewers were not representative of the total sample in such respects as employment status, television viewing, and radio listening patterns.

2. The special educational program reached 88 of the 246 respondents or approximately 36 per cent of the cross-section sample. More than three-fourths of this audience was exposed through only one medium, i.e., 9 per cent by radio, 15 per cent by television, and 55 per cent by newspapers.

3. No statistically significant differences were found between those who were exposed and those who were not exposed to the special program in terms of

employment status, age composition, or levels of education.

4. "For the most part, differences in levels of knowledge"[1] between those exposed and those not exposed were statistically significant.

5. The percentage of those who bought or served pork during the week subsequent to the special program was noticeably greater among those exposed than among those not exposed.

Of the several conclusions and implications presented, the following were included:

1. That the collected data indicated that the special consumer marketing program was highly successful. This was reflected in the size of the audience, in changes in levels of knowledge, and in consumer behavior.

2. That the St. Joseph consumer marketing information had been reaching a sizeable audience, with considerable regularity.

3. The findings further demonstrated the feasibility and the desirability of incorporating systematic evaluation into the planning and conduct of an educational program.

4. The study demonstrated a useful and practical procedure for evaluating consumer marketing programs.

5. The findings suggested many implications for future program planning.

6. Findings provided further evidence of the potential hazards involved in orienting programs exclusively for those who are most vocal in expressing their interests since the "write-in" group was unrepresentative in many respects to the total sample.

---

[1] *Ibid.*, p. 11.

Other studies listed in the Bibliography which are illustrative of the same type of study include those reported by the Federal Extension Service and by Gibbs.[1]

Welch, applying the adoption concept as a criterion of measurement, evaluated three adult education processes for disseminating trade information to Missouri restaurant operators.[2] The objectives of the study were to:

1. Evaluate two discrete processes and a combination of those processes of adult education used in an educational program for restaurant operators by employing the adoption concept;
2. Determine if individual characteristics of intra-occupational group socioeconomic status, social participation status, age, sex, education, and prior managerial experience of participants were significantly related to the adoption of recommended practices; and
3. Determine the nature and extent of the "indirect" processes that channelled knowledge and/or information to the restaurant operators in Missouri cities.

The study was conducted with the members of the Food

---

[1] Federal Extension Service, Better Informed Consumers, Circular 502 (Washington, D. C., 1955); and M. A. Gibbs, Indications of Homemakers Motivations and Knowledge about Food Buying (East Lansing: Michigan State University Agricultural Extension Service, 1959).

[2] John M. Welch, "An Evaluation of Three Adult Education Methods for Disseminating Trade Information to Missouri Restaurant Operators" (unpublished Ph.D. dissertation, Florida State University, 1961).

Service Industry in the four metropolitan areas of the state of Missouri. The first step was the conduct of a "restaurant management clinic" in each of the respective areas. Capitalizing on the motivating title for these clinics, "The Anatomy of Profit," the educational content diffused was job-related, and emphasis was given to those discrete measurable skills important to members of the Industry. The clinics were widely publicized; they were open to any member of the Industry, and the cost of attendance was kept at a minimal level in order not to exclude any potential participant. A complete roster of participants attending all sessions of each clinic was prepared.

Welch selected for testing three familiar adult education processes used in the diffusion of information: (1) an adult educational group method; (2) a mass communication method; and (3) a combination of both. The group method, identified as a restaurant management clinic, was a workshop consisting of four half-day sessions. Each session consisted of a lecture followed by a demonstration which in turn was followed by small work groups assigned problems related to the material presented previously. This clinic design had been tested previously for industry utility and for participant satisfaction by the investigator using members of the

Food Service Industry in Florida.

The mass communication method consisted of an Extension Circular which was distributed free to a sample of the universe which did not attend the clinic. This circular, specially prepared for this study, contained the same basic educational content emphasized in the clinics. The purpose of the combined method was to determine whether the impact of the two processes would produce results superior to either process alone. Therefore, the Extension Circular was mailed to one-half of the clinic participants after the completion of the clinic sessions. In the selection of educational content to be diffused, seven discrete and identifiable practices of proved utility to the Food Service Industry were emphasized both in the clinics and in the circular.

Random samples were drawn in the four urban centers in proportion to the number of food services establishments reported for the area in the 1958 census of retail trades. These four samples represented 4.12 per cent of the total number of food service establishments in the four metropolitan centers. The four area samples were divided into four subsidiary sample groups. Sample A, consisting of thirty-four, was drawn from the roster of clinic participants. This sample was randomly divided into two parts, A_1 and A_2.

To provide data for the combined method, the Extension Circular was mailed to those members of Sample A_2 after the completion of the clinics. Samples B and C, consisting of thirty-three each, were drawn from a reliable list of all establishments in the four centers from which the names of clinic participants had been deleted. The Circular was distributed by mail to members of Sample B. Those persons constituting Sample C did not participate in either of the educational processes and were used as a control group. The use of the control group provided a measure of diffusion of knowledge through indirect means.

In summary, the four sample groups used to test the relative effectiveness of the selected processes were as follows:

1. Sample A_1, consisting of seventeen Ss: Those who participated in the clinic only;
2. Sample A_2, consisting of seventeen Ss: Those who participated in the clinic and also received the Circular;
3. Sample B, consisting of thirty-three Ss: Those who received the Circular only; and
4. Sample C, consisting of thirty-three Ss: Those who neither participated in the clinic nor received the circular--the control group.

Four instruments were developed and used as follows.

1. The Survey Instrument. This interview guide was

used to collect and record all data during the personal interviews held with the 100 respondents.

2. The Intra-Occupational Group Socio-Economic Status Scale. This scale was used to determine the socio-economic status of respondents.

3. The Social Participation Scale. This scale, a modification of the Lionberger-Coughenour Social Participation Scale, was used to determine the extent of each respondent's participation in those formal organizations open to members of the restaurant industry.

4. The Degree of Adoption Scale. This scale, devised to measure the degree of adoption of the seven recommended practices by each respondent, enabled the investigator to compute three types of adoption scores for each respondent: (1) the extent to which recommended practices were in use prior to the study, therefore, attributable to prior influence (AS_p); (2) the extent to which practices were adopted as a result of the method used (AS_s); and (3) the total degree of adoption from all sources (AS_t).

Adoption scores for each respondent were computed as follows. For each of the five steps in the adoption process, a value of 0.2 was assigned. Thus, complete adoption of any one recommended practice resulted in a score of 1.0 ($5 \times 0.2 = 1.0$). Component steps within the degree of adoption concept were assigned to the source from which each resulted. For example, if it were determined that the respondent had heard about a practice from a prior source (score 0.2) and that interest in it had been aroused from that source (score 0.2), but that the application of the practice to the respondent's own business (score 0.2), the trial of it (score 0.2), and the evaluation of the trial leading to final adoption (score 0.2) resulted from information received in the study diffusion process, then the resultant scores were as follows:

AS_p = Awareness (0.2) + Interest (0.2) = 0.4

AS_s = Application (0.2) + Trial (0.2) + Adoption (0.2) = 0.6. Therefore,

$AS_t = AS_p\ (0.4) + AS_s\ (0.6) = 1.0.$

Thus, the total score for each respondent ranged from zero (total unawareness of any of the seven recommended practices) to 7.0 (total adoption of all practices recommended). All or any part of this score could be assigned to either prior influences or to the diffusion method used.

At the end of an "impact" period of approximately two months following the completion of the clinics and the distribution of the circulars, data were collected by three interviewers who used the personal interview guide. All interviews consisted of both interrogations and observations. Two procedures were used to insure the internal reliability of these interviews: (1) the interviewers worked together during the preliminary phases of the survey to develop a uniform interviewing and recording technique, and (2) a "lost record" re-interview technique was employed. Each interviewer re-interviewed a respondent previously interviewed by each of the other interviewers. Results were compared for discrepancies and for uniform scoring.

All data were statistically analyzed, and results were clearly presented in twenty-six tables. The lengthy report described and interpreted both the statistical analyses and the resulting findings, the most important of which are summarized in the following paragraph.

The individual adoption scores resulting from prior influences (AS_p) were subjected to a single analysis of variance within and between the four sample groups. Results indicated that the groups were not significantly different at the .05 level of confidence. Therefore, they were random samples of the same universe, and the AS_p scores were considered as the base line to measure changes resulting from the educational processes studied.

Findings regarding the relationship of personal characteristics of respondents to adoption of the recommended practices may be summarized as follows. Socioeconomic status was significantly related to total adoption score and to the adoption score from prior influences in all groups. Social participation score was significantly related to total adoption and to adoption from prior influences for the group receiving the circular only, but not for the clinic groups. Neither age nor managerial experiences was significantly related to adoption, but education was significantly related to total adoption score.

The degree of adoption of the recommended practices at the time of the interview was entered as the total adoption score (AS_t). The difference between this score and the score from prior influence resulted in the adoption score attributable to the study process, i.e., $AS_t - AS_p = AS_s$. The percentage of increase in adoption for each group was found to be: Group A_1, 113; Group A_2, 115; Group B, 42; and for Group C there was no percentage of increase.

A t-test was applied to determine the significance of the difference between the means of the AS_p and AS_t scores. All samples which were subjected to an educational process showed gains in adoption which were significant at the .05 level of confidence. No gain in adoption resulted in the control group.

A single-tailed t-test was computed between the AS_s scores of the sample groups representing the processes. Results revealed that the clinic method was significantly superior to the mass communication method, which in turn was significantly superior to the reliance on indirect influences for the dissemination of knowledge of the trade practices recommended. The distribution of the Extension Circular to clinic participants did not result in a significantly higher rate of adoption than that achieved by participants in the clinic alone.

The following conclusions were drawn:

1. The adult educational group method was significantly better than the communication method in changing behavior of restaurant operators in the state of Missouri.
2. Both methods were significantly better than indirect influences with the particular group studied.
3. Supplementing the educational method with the communication method did not result in any statistically significant increase in adoption over the educational method alone for restaurant operators in the state of Missouri.

The investigator was unable to identify other studies listed in the Bibliography which are illustrative of the type of study conducted by Welch.

Summary Evaluation

The preceding sections, pages 59 to 115, have been devoted to reporting seventeen studies under the category of Methods, subdivided into four types of studies. In keeping with the pattern followed in presenting the preceding categories, a general summary evalution of the studies in

Category III: Methods, will conclude the discussion.

Of the seventeen studies categorized as Methods, nine were descriptive in nature, and eight studies either utilized a control group against which treatment effects were compared or compared two or more groups for which different treatments had been applied. First, a summary evaluation is made of the nine descriptive studies after which the remaining eight studies will be evaluated.

The nine descriptive studies included seven in which results of one treatment were assessed in relation to the participants as a single undifferentiated group. The seven studies were conducted by Norby, Lintner, Clem, Holder, Jennings, Matott, and Reuter.[1] Two of the nine studies were those in which results of one treatment were assessed in relation to differentiated groups of participants. These two studies were conducted by Wilson and Moe and by Guarino.[2]

The sample used in each of the nine studies may be considered adequate in terms of representing the number of participants involved in each program. Random samples were used by Wilson and Moe (25 per cent), by Jennings (approximately 10 per cent), and by Lindner (4.3 per cent). Norby and Reuter obtained data on all participants. With only one exception, the sample used by Matott represented the total population, and Guarino obtained data from 118 of the 124 dairy farmers who represented the population studied. In terms of reported populations in the areas represented, Holder's sample was adequate, but some question remains as to the precise method used to select respondents who were non-club members. Although Clem limited her study to the ten counties in which agents had been employed for a period of at least three years, she undertook to evaluate the present status of the craft phase of the Tennessee Agricultural Extension Service Program in Upper East Tennessee which included twenty-one counties. Based on her historical interpretation of the crafts in East Tennessee, one may assume that the ten counties were fairly well representative of the twenty-one counties.

---

[1] Supra: Norby, p. 60; Lintner, p. 62; Clem, p. 65; Holder, p. 66; Jennings, p. 66; Matott, p. 67; Reuter, p. 68.

[2] Supra: Wilson and Moe, p. 85; Guarino, p. 87.

All data were obtained through the use of a questionnaire, or a series of questionnaires, frequently identified as an interview schedule. Reports indicated that questionnaires had been pretested by Norby, Lintner, Matott, Wilson and Moe, and Guarino. This pretesting involved a trial run with a few participants from the same population which in turn enabled the investigators either to add, delete, restate, or otherwise clarify intended meanings of items. The reports gave no information on efforts to establish validity, reliability or discriminating power of questionnaire items. Evidently the questionnaires used by Clem, Holder, Jennings, and Reuter had not been pretested since such was not reported.

In the studies conducted by Norby, Holder, Jennings, Matott, Wilson and Moe, and Guarino, questionnaires were used during interviews with respondents and Lintner, Clem, and Reuter mailed their questionnaires. Usually, researchers consider that more valid information is obtained when an instrument is used during an interview than when it is mailed. This depends, of course, on the instrument and on the competencies of the interviewers. It seems probable that the interviewers used by Wilson and Moe were more competent than were those employed in the other studies since it was reported that they were trained interviewers and since the investigators were staff members of the Federal Extension Service. The fact that the study reported by Norby was conducted in connection with an over-all research project designed to investigate the porcesses and values of extension program planning at the county level which was carried on at the National Agricultural Extension Center at Madison, Wisconsin, makes it reasonable to assume that the interviewers had received considerable training in the techniques of interviewing. For the remaining four studies in which interviews were conducted, all investigators were employed by the Cooperative Extension Service. Holder used home demonstration club members to assist her. Jennings reported that the interviews were conducted by eight women, all of whom were employed by the Cooperative Extension Service. Matott and Guarino each was assisted by one other county agent. Consequently, it may be assumed that each of the investigators had previously received some training in interviewing techniques since the Cooperative Extension Service has used this method extensively to obtain data.

It must be recognized that with the exception of a part of the studies conducted by Norby, Clem, and Matott, the research method employed for each of the nine studies was one in which judgments of the participants were obtained. In the absence of other data, the researchers were forced to accept the statements of respondents at face value. Norby reinforced this type of data with observations and tape recordings. One phase of Clem's study evaluated the quality of the craft work. The composite judgment of 120 individuals who were competent to judge craft work because of their unusually high scores on a standardized test administered for that purpose is an acceptable evaluation method. In addition to expressed opinions by the subjects, Matott analyzed farm records kept by the participants which he considered objective evidence of the changes made as a result of the program.

All investigators, except Norby, reported data only in terms of frequencies and percentages. No attempts were made to establish the significance of any of the findings as a basis for generalizing the results. Since no control groups were used, it is hazardous to infer that the changes reported in each case were the results of the several programs under study and not of other influences operating in the general populations. With minor exceptions the level of evaluation was simply that of reporting results. The methodology used by Norby may be considered a pioneering effort inasmuch as he stated that no method for evaluating a program planning process had been developed previously. The procedure he used to identify principles and to deduce the conditions embodied within the context which he supported theoretically by statements of guidelines from the professional literature and from generalizations from research is an acceptable research study within itself. However, Norby chose to use the conditions as criteria to determine the effectiveness of one county extension program planning process. Thus, the procedure which he finally developed for a program planning process based on this experience is a contribution to practice since the suggested guide may be applied in other situations.

All of the nine studies contribute to practice but their value is limited by the fact that in each case the information obtained applied to the programs specified and the basis for generalization to other programs is inadequate.

The research methodology employed for the remaining eight studies represents a higher level of research than did the research methods employed in the preceding nine studies. The use of a control group against which treatment effects were compared and the comparison of two or more groups to which a different treatment was applied represents a significant improvement over the simple design which reports changes or progress as judged by participants.

The eight studies include four in which a control group was utilized to compare effects of a single treatment. Two of the four studies were conducted by Alexander and Longest and by Holik and Moore who assessed results of one treatment in relation to a single undifferentiated group of participants.[1] The two other studies were conducted by Rogers and Havens and by Lacy who assessed results of one treatment in relation to differentiated groups of participants.[2]

In the remaining four of the eight studies comparisons were made of two or more groups, each of which had been subjected to a different treatment. Two of these four studies were conducted by Marsh and Hyman and by Robert who assessed the results in relation to an undifferentiated group of participants.[3] The remaining two studies were conducted by Porter and Wilson and by Welch who assessed results of three treatments in relation to differentiated groups of participants.[4] Welch also used a control group in addition to the three treatment groups to provide a measure of knowledge through indirect means.

The samples which were utilized in six of the eight studies may be considered adequate. These six studies were those reported by Holik and Moore, Rogers and Havens, Lacy, Marsh and Hyman, Porter and Wilson, and Welch. The two remaining studies, those conducted by Alexander and Longest and by Robert, reported that samples were not representative of the population. However, in most studies there were sampling limitations which must be recognized. These limitations were of three types; namely, those samples which may have introduced bias because of the way in which they were selected, those samples on which the reports gave insufficient information, and those participant samples which were poorly matched with nonparticipant samples. Sampling limitations are reported for each of the eight studies in the subsequent paragraphs.

---

[1] Supra: Alexander and Longest, p. 70; Holik and Moore, p. 82.

[2] Supra: Rogers and Havens, p. 88; Lacy, p. 93.

[3] Supra: Marsh and Hyman, p. 96; Robert, p. 100.

[4] Supra: Porter and Wilson, p. 102; Welch, p. 107.

Holik and Moore used two samples of subjects who were matched only on the basis of similar cropland acreage and of location of their farms. The investigators pointed out that any inferences or generalizations must be made with these limitations in mind.

Rogers and Havens utilized random area samples in both the treatment and control counties. However, the farmers who were interviewed twice represented a comparatively low percentage of the population. The investigators did not report the number of farmers whom they failed to contact in three visits. Thus, it must be recognized that some bias may have been introduced. The unidentified number of farmers who were not included were not omitted by chance but because of the operation of unknown factors which made them unavailable.

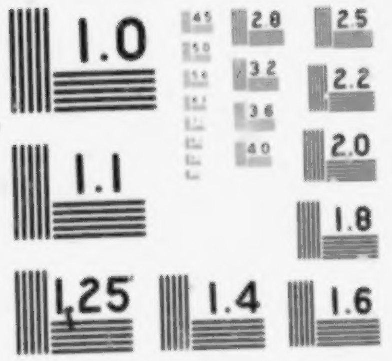
Lacy, because of the purpose of his study, was forced to use the twenty-eight members of the program planning committee as the experimental group. Members of the control group were matched approximately in terms of age, sex, occupation, and education with those of the experimental group. The investigator pointed out that neither group, from a statistical analysis viewpoint, could be considered random samples. Thus, conclusions were limited to the group studied.

Marsh and Hyman carefully controlled the selection of their random sample which was stratified by sex. One may assume that it was adequate; however, the researchers failed to report the number of eligible county agents in the state from which the sample was drawn or the percentage of the population which the sample represented. Moreover, in terms of the stated hypotheses, their justification for stratifying by sex may be questioned. The report gave no indication

3 OF 5

ED

003520


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963

that any comparison of the differences in gains in test scores was made between male and female subjects. Since much current educational research is directed toward sex differences in learning, the investigators may have missed an opportunity to contribute to the growing body of knowledge on sex differences in education. It may be assumed that the investigators were more interested in providing in-service preparation experiences for agents while comparing the two methods than in conducting research which contributes to knowledge. This is their prerogative. Nonetheless, readers of this report may ask, "What difference would it have made in the findings if the group of new agents had been predominately women?"

Porter and Wilson used a cross-section sample of 246 persons drawn at random from the St. Joseph telephone directory. This sample was representative of the population in terms of age distribution if it can be assumed that the female population of St. Joseph was similar to that of the state according to the 1960 census population. A table in the Appendix presented the age composition of sample homemakers and the total Missouri female population of twenty years and over. Also, the investigators reported in the Appendix that additional comparative data on education and employment status would be used to validate the sample as soon as the information was available. Moreover, the investigators reported neither the number of the people who requested the skewers nor the number included in the sample drawn from this "write-in" group.

Welch used random samples which represented 4.12 per cent of the total number of food service establishments in the four metropolitan centers in Missouri. However, the samples should be viewed as two not strictly equivalent sub-samples. Sub-sample A was made up of individuals who voluntarily sought knowledge and Sub-samples B and C were drawn at random from the remainder of the universe of restaurant operators. To an unknown extent, the members of Sample A may have been more favorably predisposed to the adoption of new practices than were those of Samples B and C. Hence, the reported differences favoring Sample A may be a somewhat inflated estimate of treatment effects. In a field in which participation is voluntary, sampling of instructed and of non-instructed populations may be expected to be biased to some degree. Subject to this limitation, the samples employed in this study were adequate.

Alexander and Longest encountered difficulties with their samples and the report of the study described in detail the procedures which were employed to deal with these difficulties. The original samples were adequate, and results of statistical analysis indicated that the participant sample was fairly representative of the universe of participants. However, because of attrition in both groups during the four-year period, the final samples which were used to measure change in the stated phenomena included eighty-seven matched participants and nonparticipants. The eighty-seven participants were not representative of the universe of participants. The writer agrees with the investigators that the increased control attained by using pair-matched subjects in an experimental and control group was preferable to using participants who were poorly matched with nonparticipants. Since the investigators pointed out that comparisons of the eighty-seven matched participants and nonparticipants would result in the best estimate of the effectiveness of the program, it is the judgment of the writer that this decision enhanced the quality of the study. Nonetheless, the final samples were inadequate for the stated purposes of the study.

Robert acknowledged that he made no attempt to be selective in establishing the three groups. He also pointed out that only limited controls could be placed on establishing the group of farmers who attended a neighborhood meeting.

Therefore, since the samples could not be considered representative of the population, they were inadequate for stated purposes of the research.

The treatment(s) which were employed were adequately described for those studies conducted by Alexander and Longest, Rogers and Havens, Lacy, Robert, Porter and Wilson, and Welch. The report by Marsh and Hyman lacked precision in the description of methods compared, particularly the workshop. Because of its simplicity and of the increasingly growing use of the self-instruction method, it may be considered adequate simply to report that self-instruction was used. However, in the case of the workshop, this is a crucial omission inasmuch as the literature abounds with variations of organizational and instructional patterns; of teaching techniques and materials; of time schedules in terms of duration and in terms of time allotted to total group, small group, and individual activities; and of approaches made to identify problems or subject-matter content to be considered or presented. Conducting a workshop for which the subject-matter content was selected previously by those in charge, and designating a two-day time schedule, are major modifications of the workshop method as initially conceived by its originators in the mid-thirties.

The workshop method employed by Welch, identified as a restaurant management clinic, also was a major modification of the original concept of the workshop, per se. However, he described in detail the manner in which it was conducted, and pointed out that he had previously tested this "clinic design" for industry utility and for participant satisfaction with restaurant managers. Consequently, in terms of methods which were compared, other researchers could replicate Welch's study but they would have to consult Marsh and Hyman before they could replicate their study. Holik and Moore did not describe the Balanced Farming Program; however, it must be noted that a reader of the report becomes somewhat informed of its major purposes from the lengthy descriptive interpretations of the findings which are presented.

Even though treatments were adequately described in most cases, the statements of objectives of given programs left much to be desired. For the most part, objectives of programs were stated in generalities. The first requisite in evaluating an adult education program is to state precisely the objectives of the program. Secondly, the objectives must be translated into specific outcomes which are representative of the expected behavior of participants when objectives have been achieved. Only then are the outcomes measurable, and they must be measured in terms of the movement from a base line. In other words, before behavioral change either in attitudes, understandings, knowledges, skills, or practices can be determined, one must make an assessment of the extent to which participants possess such attributes before a specified treatment is applied. This assessment provides the base line.

The statement of an objective by Marsh and Hyman represents a more precisely stated objective for the program than was stated in most reports. However, in this study, as in most of the other studies, other objectives of the given program either were expressed or implied. Marsh and Hyman reported: "The primary objective of the training program was to increase the agents' knowledge of the use of radio in Extension teaching and thus hopefully to increase the competence of the agents in the use of this medium."[1] In two studies the objective(s) of the program were stated precisely and the expected behavioral outcomes were at least inferred for those participants who achieved the objective(s). Robert had one precise objective: To get farmers to check their rice fields for symptoms of the rice disease--Hoja Blanca. Thus, the expected behavioral outcome was obvious. Porter and Wilson stated specifically ten basic concepts to which the special program on pork was geared. These concepts reflected the behavioral outcomes expected of participants who developed the concepts. However, the study did not ascertain the extent to which exposed respondents possessed these concepts at the time the special program was begun; hence, outcomes were not measured in terms of the degree of movement from a base line.

---

[1] Marsh and Hyman, op. cit., p. 7.

Welch's study is the only one of the twenty-two investigations conducted under the auspices of the Cooperative Extension Service which applied all the proposed requisites of evaluation. The specified objective of the program was to secure the adoption of seven discrete, identifiable, measurable practices which were recommended to restaurant operators. The explanation of these practices and the emphases which were placed on their use clearly indicated the expected behavior of restaurant operators when they adopted the practices. A copy of the Extension Circular which presented this information was included in Welch's report. The same educational content was emphasized in the restaurant management clinics. In order to measure the behavioral outcomes (the extent to which participants had adopted the recommended practices) resulting from the treatments, Welch established a base line for each respondent. This was done through the use of the Degree of Adoption Scale which was specially designed for the study to measure the extent of adoption of the seven recommended practices by each respondent. This scale enabled the investigator to compute a score of the degree to which each of the seven recommended practices were in use by the respondent prior to the treatment (AS_p). This score was the base line. Hence, the total adoption score after treatment (AS_t) minus the AS_p score equalled the score which had resulted from the treatment (AS_s). Thus, the behavioral outcomes were measured in terms of the participant's movement from a base line.

Despite the lack of precise statements of objectives for a given program and the failure to translate the objectives into expected behavioral outcomes which would be indicative that objectives were attained, efforts were made to determine the participants' movement from a base line as a result of a given program. In six of the eight studies a base line was established and the extent of change in participants from the base line resulting from treatments were determined. Besides the study reported by Welch, other such studies are those by Alexander and Longest, Holik and Moore, Rogers and Havens, Lacy, and Marsh and Hyman. The criterion measures used in these studies indicated that more precise objectives of programs were perceived than were stated by the researchers. The quality and type of these measures are discussed in the following paragraphs which report the ways in which data were obtained.

With the exception of the study by Marsh and Hyman, the data for the eight studies were obtained by interviews with respondents. The reports of studies by Alexander and Longest, Lacy, Porter and Wilson, and Welch stated that interview schedules had been pretested. Rogers and Havens reported that most test items included in their schedule had been used in previous research, but they did not imply that other sections had been pretested. In the reports of the studies by Holik and Moore and Robert no reference was made to the pretesting of schedules. Robert, Porter and Wilson, and Welch conducted interviews only once, following the treatment; all other investigators conducted interviews both before and after treatments.

All interviewers had received at least some training in interviewing techniques and most of the interviewers may be considered competent because of their experience in conducting sociological research. Porter and Wilson used volunteers who were trained by the Extension staff. These volunteers used a carefully-constructed schedule which also included statements relating to appropriate social amenities for telephone interviewing. It is reasonable to assume that all telephone interviews were conducted uniformly and politely. Welch reported that observations were made by the interviewers during their visits to conduct the interviews with restaurant operators. Since the seven recommended practices were "identifiable," the data obtained by Welch from respondents and from observations which were made at the same time are likely to be more valid than those data obtained only from respondents' statements.

Of the eight studies, special scales or tests designed to measure change in attitudes, in knowledge or in adoption practices were used in six of the investigations. The six were those studies by Porter and Wilson, Lacy, Marsh and Hyman, Alexander and Longest, Rogers and Havens, and Welch. Porter and Wilson designed ten questions which related specifically to the special information which was disseminated about pork. Five of the questions related to the ten concepts and five related to pork consumption patterns. It may be assumed that these ten items possessed face validity, but no information as to the quality of tests was reported. Lacy ascertained attitudes of subjects toward the Cooperative Extension Service by using a Likert-type scale of summated scores. To measure changes in knowledge, he used a series of questions which were related to those specific areas of knowledge of the Cooperative Extension Service judged to be emphasized during the program planning process. These tests were carefully constructed but the report gave no information pertaining to efforts to establish validity or reliability of items. Lacy pointed out, however, that the measures may not have been sufficiently refined to detect minor changes in attitudes and in knowledge which might have been attributed to the treatment employed. It may be assumed that some effort had been made to establish the validity and reliability of scales used by Marsh and Hyman since they had been developed previously by Tait and Hatch and since Tait used the tests in a similar study of which the Marsh and Hyman study was a replication.[1] However,

---

[1] Elton B. Tait and J. Cordell Hatch, "Make Radio Work for You" (Pennsylvania State University Agricultural Extension Service) and Elton B. Tait, A Comparison of Self-Instruction and a Workshop as Methods of Training Newly Hired County Extension Personnel in the Subject of Radio Broadcasting, Extension Studies No. 15 (University Park: Pennsylvania State University Agricultural Extension Service, 1962), cited by Marsh and Hyman, op. cit., pp. 4-5.

Marsh and Hyman did not report any analyses of the tests.

As indicated in the analyses of the three remaining studies, those reported by Alexander and Longest, Rogers and Havens, and Welch, considerable efforts had been made to establish the validity and reliability of instruments used. Alexander and Longest assumed face validity of their tests because of the competencies of the subject-matter specialists who constructed the original items. They employed statistical analyses to establish reliability and discriminatory power of items; acknowledged that some tests contained items that were low in reliability and discriminatory power; and presented their rationale justifying their decision to use the tests. Thus, as much as possible was reported about the quality of the measures. Nevertheless, it should be recognized that such tests might have been of higher quality had the investigators tested items before the administration of the initial schedule and thus have made possible the inclusion of better items.

Rogers and Havens also employed several statistical analyses to establish validity and reliability for the attitude scale. Since the judges' agreement on the final scoring was found to be relatively high, it may be assumed that this scale was of high quality for this particular case.

However, the basic premise of Rogers and Havens that the use of projective techniques avoided bias induced by attempting to give socially acceptable answers is open to considerable question. It may be assumed that their knowledge test possessed face validity.

The three measures used by Welch had been slightly modified from scales used in previous sociological research. It may be assumed that the respective authors who developed the scales had established validity and reliability. The scales are probably of high quality since they are used frequently in similar types of research.

In six of the eight studies data were statistically analyzed. Differences found between the first and last assessment of participants were tested for significance by using either the t-test or the F-test, and confidence levels were established at the .05 and higher levels.

All conclusions for the eight studies appear to be sound and warranted. In the two studies in which the samples were inadequate, those by Alexander and Longest and Robert, the investigators made no claim of drawing conclusions as such.

All studies contribute to practice. Technically, the findings of the foregoing eight studies may be generalized only to the populations represented by samples. This statement is of course true of all sampling studies. It is the responsibility of the investigator to report his population characteristics carefully and fully in order that the reader may estimate the applicability of the findings to other populations in which the reader is interested. Most of the research by Cooperative Extension Service has reported such characteristics of subjects as sex, age, education, employment status, size of farm, and others.

University Extension

Of the evaluation research conducted under the auspices of University Extension, five studies were identified which relate to the occupational function of Adult Education. These five studies are not readily classifiable, but they may be ordered in terms of their sophistication. One study should be placed in the lowest rank. It was essentially an *ex post facto* evaluation of a training program. Three studies having better designs and utilizing more sophisticated analytical techniques dealt with the comparative outcomes of three specific campus vs. extension classes. These three studies may be considered to be in the second rank of University Extension studies. The fifth study was essentially an empirical test of a theoretical evaluative technique and should be considered in a rank by itself. Accordingly, these five studies are analyzed sequentially in terms of their order of sophistication from the lowest to the highest rank. A summary evaluation of the five studies completes this section.

Murphy evaluated a program of many years standing which was conducted by the University of California, Berkeley, for executive personnel of the Allied Airlines.[1] The program, limited in content to economics and leadership training, was a four year program which incorporated four one-week residence seminars during each of the four summers, two intervening correspondence study courses, and one year devoted to a guided reading program. Accordingly, a graduate of this program was required to complete the following six courses in order and by the method indicated below.

Psychology of Human Behavior	Residence
Psychology of Communication	Correspondence
Effective Leadership	Residence
Economics and Business Theories	Correspondence
Business Fluctuations and Forecasts	Residence
Advanced Human Relations and Leadership	Residence[2]

---

[1] Thomas F. Murphy, "The University's Role in Management Development" (unpublished M.B.A. thesis, University of California, Berkeley, 1959).

[2] *Ibid*., p. 27.

The guided reading program came between the third and fourth residence sessions. A typical listing of the books which was given consisted of nine carefully selected titles.

To evaluate the program Murphy attempted to measure:

1. Degree of acceptance of training by older executives; and
2. Perceived value received for the company's training dollar.

Specifically, he sought answers to two questions:

1. Do mature executives react favorably to training programs, or do they consider them a "necessary evil imposed by management"?
2. Is the value received by the company for its training dollar real or imagined?[1]

A random sample of 50 was drawn from the total of 134 executives who had completed the management training program. Two instruments were used for measurement.

Murphy constructed a questionnaire which was designed to provide, in addition to general background information, forced choice preference statements which reflected preference for elements of the program format and course content. It was pointed out that the questionnaire was used in an effort to obtain a subjective evaluation of the format and of the methodology employed in this program and to determine the degree of acceptance of this and subsequent training programs. A copy was included in the Appendix of the report. This questionnaire was mailed to the forty-eight randomly selected graduates. Two of the original graduates were no longer employed by the Company. Forty-one, or 85 per cent, of the questionnaires were returned. Data were organized by year of graduation, but since no significant differences were found, the forty-one respondents were considered as a group.

---

[1] *Ibid.*, p. 29.

The test, "How Supervise?" by Quentin W. File and H. H. Remmers was used. Murphy pointed out that this instrument was used because of its reliability as a measure of a person's understanding of the general aspects of supervision and because it permitted expression of attitudes toward supervisory practices and toward company policies. To measure the existence of the "ripple effect" in the Allied Company, this test was administered to those of the same sample, and to a control group of forty-six Allied supervisors who were of comparable job level, age, and educational background, but who had not participated in the Management Training. However, these forty-six supervisors had been selected to begin the program in the forthcoming summer. The "ripple effect" referred to the educational value which had spread outward from and beyond those who had participated in the program and hence influenced the nature of the broad management climate of the company. Personnel of the Education and Training Department of the Allied Company had commented favorably on the "ripple effect" of the program. Forty-one tests, or 85 per cent, were returned from each of the sample groups.

These data were analyzed and the report presented a lengthy discussion of the findings. Briefly, they may be summarized as follows:

1. The typical graduate was perfectly happy with the format of the program, but if changes were necessary he preferred a two weeks residence in each of two summers with two intervening correspondence courses.
2. The graduate preferred the residence session over the home study method regardless of the subject matter to be learned.
3. He felt strongly that a post-graduate course should be provided about two years after completion of the Management Training Program, but was completely indifferent as to whether new subjects or more advanced work in human relations and economics should be undertaken.
4. Each graduate felt that the Program has been extremely valuable to him in his job.
5. Upon examination of the test scores, no statistically significant difference existed between the scores achieved by the two groups.

After obtaining these results the investigator attempted to explain the outcomes on the bases: (1) of the unreliability of the instruments, (2) of the low validity of the instruments, and (3) of the power of the "ripple effect." On the basis of the test authors' data, he concluded that the instrument was reliable. On a logical basis, he concluded that the test may have been only minimally valid. He stated that the assumption of the operation of the ripple effect might be expected to have more validity than either of the preceding assumptions. A fourth possibility, namely, that "the training experience may create change in individual behavior without appreciably altering the group's index of human relations skills,"[1] was cast as a null hypothesis and tested by the chi-square technique. He found that this was significant at a confidence level of .20; hence, the data were interpreted as indicating a trend but because of the low confidence level, the findings were not accepted as being statistically significant.

In presenting the conclusions, Murphy pointed out that the responses of the forty-one executives were "overwhelmingly favorable" to the program. Thirty-four, or almost 83 per cent, wrote "glowingly of the value of the program in response to the request to list specific instances in which the training had been helpful."[1] He concluded that "nothing in the data, which were furnished anonymously and which could be considered as valid as any attitude survey could be indicated any support for the supposition that executives of Allied Airlines felt anything but pleased by their selection to engage in this training program."[2] He also pointed out that the "measurement of the value received by the Company was not as clearly brought out by the data; however, there was a rather significant trend in the direction of the original hypothesis."[3]

---

[1] Ibid., p. 51.

The following three studies were considered to rank in the second place in terms of the order of research sophistication. Essentially, their research designs are similar. Each study is reported, however, because of differences in criterion measures used or in procedures employed to analyze data.

McCormick compared the performance of two groups of students completing course work in both the extension program and the resident program offered by the University of

---

[1] *Ibid.*, p. 58. [2] *Ibid.* [3] *Ibid.*, p. 59.

Virginia.[1] The two groups were: (1) the "before" group, consisting of twenty-two students who had completed their extension work before entering the resident program; and (2) the "after" group, consisting of fourteen students who completed their extension work after they had completed at least one year of their resident work. The average grade values were determined by dividing the total number of grade points earned by the total number of semester hours completed. In order to reflect the quality of the student performance, the grade points were weighted by multiplying the total semester hours of credit earned by each student by the average grade for the total hours taken. He found that the grade averages for the extension courses were higher than those for the resident courses in both groups. He also examined the performance record of twenty-five randomly selected resident students who had completed seven semesters of work, and found a consistent and accelerating decrease in semester grades during the first five semesters, the approximate academic period under study. The investigator stated, in effect, that on the basis of this analysis, the obtained difference between extension and residence grades should be reduced by ten per cent to obtain a better estimate of the difference favoring extension work. He suggested that a comparison of the performance records of students in the two programs should be made periodically and that when sufficient information is available to be significant, this comparison should indicate the effectiveness of the academic administration of the extension program.

---

[1] Fred C. McCormick, "A Comparative Study of Engineering Extension and Resident Programs," Adult Education, Vol. 11 (1961), pp. 106-109.

Costin and Johnston compared the achievement of extension students with that of campus students in an introductory psychology course by studying three campus and three extension classes over a period of three consecutive semesters, with one of each being taught each semester.[1] The same instructor taught all classes; each class met for the same number of minutes each week; and the course content and organization, examinations, and grading procedures were the same for each class. A teacher-prepared objective examination was administered following the completion of each of the five major topices studied. To obtain an approximate measure of verbal learning ability, Part I of the SCAT, Form 1-a, was administered to all classes.

For purposes of comparing achievement, seventy pairs of male students were matched on the basis of identical verbal ability scores. All data were analyzed statistically. Differences in mean scores and standard deviations were tested for significance at the .01 level of confidence. No significant differences were found between the two groups for two of the major topics studied, while campus students scored significantly higher than did extension students on three topics. However, the differences between mean scores of the two groups while significant were rather small.

---

[1] Frank Costin and Robert L. Johnston, "A Study of Achievement in an Introductory Psychology Course," Adult Education, Vol. 12 (1962), pp. 120-126.

Three other analyses were made to determine the relationships that existed between verbal ability and achievement in examinations.

1. When the achievement of high ability extension students was compared with that of low ability extension students, there were no significant differences for any of the five course examinations. A similar comparison for campus students showed a significantly higher mean examination score for the high ability students on four topics, with no significant difference on one topic.

2. When achievement of campus students was compared with achievement of extension students according to level of academic ability, high ability campus students did significantly better on examination than did the high ability extension students for four of the topics, and no difference was found for one topic. No significant differences were found between low ability campus students and low ability extension students.

3. When product-moment correlations were made between course achievement and scholastic ability, there

> were significant correlations for the campus students on all five topics, while for the extension students there were significant correlations in only two instances.

It was concluded that it is reasonable to expect extension students taking an introductory psychology course to achieve as well as campus students taking the same course. However, it was pointed out that "the results of the present study have immediate validity only for the particular course, instructor, and populations which were involved."[1]

Anikeeff compared the scholastic achievement of extension students with campus students in a personnel management course.[2] He sought an objective answer to the question: "Should students receive regular college credit for work completed under an off-campus extension program?"[3] The design of the study was a replicated two group (extension class and campus class) experiment. The treatment was carefully controlled in that the text and lecture material were uniformly presented by the same instructor. The measuring instruments were two objective, four-choice tests, each of which contained fifty items. They had been pretested and refined by maximizing the internal consistency of each test. Subjects were carefully matched on basis of pretest scores. The objectives of the course were defined in terms of cognitive learning only, and the total evaluation was directed specifically toward the achievement of these objectives. These objectives were operationally defined as scores on the tests.

---

[1] Ibid., p. 126.

[2] Alexis M. Anikeeff, "Scholastic Achievement of Extension and Regular College Students," Journal of Applied Psychology, Vol. 38 (1954), pp. 171-173.

[3] Ibid., p. 171.

Six comparisons of educational achievement between groups were made for each of the two examinations which were administered: (1) extension students before studying vs. extension students after studying, (2) day students before studying vs. day students after studying, (3) extension students before studying vs. day students before studying, (4) extension students before studying vs. day students after studying, (5) extension students after studying vs. day students before studying, and (6) extension students after studying vs. day students after studying. Anikeeff found that:

> 1. The arithmetic means of regularly enrolled students on pretests did not differ significantly from the posttest arithmetic means of extension students on identical examinations.
> 2. The posttest mean of extension students on the first

> examination did not differ significantly from a mean which could be obtained by random guessing. On the second examination, the posttest mean of extension students differed at the 5% level of confidence from the mean which could be obtained by random guessing.[1]

In discussing the results, Anikeeff questioned the advisability of granting college credit for work performed in evening off-campus extension courses. After pointing out several factors which may have contributed to the difference in achievement scores, he stated: "However, despite the reasons for the differences between the achievement of extension and regularly enrolled students, if further studies support results found in this investigation, the case for granting college credit for extension work will be severely challenged."[2]

Of the five studies identified which have been conducted under the auspices of University Extension, the following study is considered to be the highest order of research sophistication represented. Because it is an empirical test of a theoretical evaluative technique, it is placed in the highest rank.

Schwartz applied the logic of semantic

---

[1] Ibid., p. 173.

[2] Ibid.

differentiation to evaluate an adult education program.[1] The primary concern was to determine: (1) the effectiveness of a particular program of instruction, (2) the effectiveness of university professors as instructors in an industrial setting, and (3) the effectiveness of basic teaching techniques. The research design was that of a field study with a pre-and-post examination of two groups of subjects who were enrolled in the same course.

The initial problem of the study was to define a criterion measure for teaching effectiveness. Through a comprehensive review of social psychological literature, Schwartz abstracted the effectiveness criterion which he defined as "the change in meaning attributed by the student to his perceptions, vis á vis the instructor."[2] According to this criterion, if no change in meaning occurred during the instructional period, either (1) the instructor had not been effective, or (2) there was no need for instruction in the first place; i.e., the participant not only possessed the relevant information but he espoused the same attitudes and values as did the instructor. To measure "meaning" and changes in "meaning," a semantic differential instrument was constructed. This instrument was designed according to the logic of semantic differentiation and according to the form described by Osgood, Suci, and Tannenbaum.[1]

---

[1] Michael Schwartz, "Adult Education in Industry: A Program Evaluation" (unpublished M.A. thesis, University of Illinois, 1959).

[2] *Ibid*., p. 24.

Ten concepts, or stimulus words, representing those that would be most meaningful to the participants, were chosen. Eight were selected directly from the course outline: (1) leadership, (2) communication, (3) supervision, (4) group, (5) manager, (6) authority, (7) organization, and (8) subordinate. The remaining two concepts were: (9) yourself, designed as a measure of self-concept, and (10) this class, designed as a measure of perception of the class. Each concept appeared at the top of a separate page to be rated against eighteen different scales, each of which contained seven points. The scales were representative of the three dimensions of meaning, i.e., evaluation, potency, and activity. For the evaluative factor, the polar adjectives were: (1) fair--unfair, (2) valuable--worthless, (3) clear--hazy, (4) agreeable--disagreeable, (5) clean--dirty, and

---

[1] C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, The Measurement of Meaning (Urbana: University of Illinois Press, 1957), cited by Schwartz, ibid., pp. 11-32.

(6) good--bad; for the potency factor or dimension, the polar terms were: (7) strong--weak, (8) large--small, (9) shallow--deep, (10) hard--soft, (11) independent--dependent, and (12) severe--lenient; and for the activity factor, they were: (13) fast--slow, (14) active--passive, (15) hot--cold, (16) sharp--dull, (17) high--low, and (18) difficult--easy.

Because of the possibility that some subjects might develop a response-set to these scales, they were randomized; i.e., all scales representing a single factor were not placed consecutively. Furthermore, the polar words were alternated. The following two scales are illustrative.

Fair: ___ : ___ : ___ : ___ : ___ : ___ : ___ : Unfair

Weak: ___ : ___ : ___ : ___ : ___ : ___ : ___ : Strong

Subjects were to check each scale by using the quantifiers "extremely," "quite," "slightly," and "neither" as follows: "extremely" in spaces (1) and (7), "quite" in spaces (2) and (6), "slightly" in spaces (3) and (5), and "neither" (or does not apply) in space (4). To illustrate:

A: (1) : (2) : (3) : (4) : (5) : (6) : (7) : B

Scales indicated both direction and intensity. For example, if the subject placed his check mark in either of the spaces (1), (2), or (3), the direction of his response was toward polar term A. Likewise, if he placed the check in either of the spaces (5), (6), or (7), his response was in the direction of polar term B. Intensity was noted in the same manner; the assumption was that the closer to A or to B, the greater was the intensity. Thus, in summary, the semantic differential instrument consisted of ten concepts, or stimulus words, to which each of the eighteen seven-point scales were to be applied.

Schwartz proposed eight crucial questions for which he sought answers from data to be obtained from this instrument. These questions, with interpretative comments, were:

a. Does perceptual change for the students occur over time? That is, with reference to the selected course-relevant stimuli which are presented via the semantic differential, do the students show significant changes in the meanings which they attribute to their perceptions?
b. What, if any, relationship exists between the two groups before instruction begins? The determination of comparability between groups before instruction begins may have possible implications regarding various teaching methods.
c. What, if any, relationship exists between the two groups after instruction has ended? What changes, if any, have occurred over time between the two groups?
d. What relationship, if any, exists between groups with reference to variability over time? Significant change in variability of response over time may have implications for teaching methods.
e. Over time, what has happened to the variability of responses for each group? This question when

answered would provide supporting evidence for an answer to (d).

f. What is the relationship, if any, between instructor and students as to the way in which they perceive selected course relevant stimuli? The answer to this question may provide an answer to the problem of determining training needs. If the instructor differs greatly in his perceptions from the group, then one might assume that there is need for training.

g. At the end of instruction, what is the relationship between instructor and students? That is, has the degree of homogeneity of response between instructor and students increased over time, and is the instructor primarily responsible for this increase or are the students?

h. Are certain concepts or dimensions of meaning found to be more variable over time? This perhaps may provide an answer to the question of "area of effect." That is, it may be possible that the instructor is more effective in the teaching of one area than another.[1]

The "particular" program was a course entitled "Managers and Their Relationships" which was part of the regular program sponsored by the General Extension Division of the University of Illinois. Participants were first-line supervisors of the state power company who were divided into two groups on the basis of the respective area in which they worked. According to the Director of Training for the company there was little variance in age, education, and length of service between participants in the groups. One group was identified as the Homeville Group and it was taught by a university professor who used primarily the lecture technique. The other group, identified as the Doe City Group, was taught by a different university professor who employed primarily the group discussion technique. The course was designed for thirty-two hours of instruction pertaining to the management function, and the same course content was used for each group. The semantic differential instrument was administered to both groups and to both instructors before and after instruction.

---

[1] *Ibid*., pp. 25-26.

The data were analyzed statistically in terms of the proposed questions. Critical ratios, F ratios, and the sign test were employed in four types of analyses, i.e., groups were analyzed for change in meaning over time, and for change in variability of meaning over time as follows:

G_{1_1} vs. G_{2_1} (group one at time one vs. group two at time one);

G_{1_2} vs. G_{2_2} (group one at time two vs. group two at time two);

G_{1_1} vs. G_{1_2} (group one at time one vs. group one at time two);

G_{2_1} vs. G_{2_2} (group two at time one vs. group two at time two).

Two other types of analyses were made: (1) between instructors' scores and the scores of their respective groups both before and after instruction, and (2) between instructors' scores at time one and time two, and between scores of the same instructor at time one and time two.

The findings may be summarized briefly as follows:

1. No significant differences were found between the two groups before instruction nor after instruction. Thus, in terms of mean scale scores for all concepts, the two groups were initially comparable and retained this comparability over time.

2. In terms of mean scale scores, neither the Homeville Group nor the Doe City Group changed significantly over time. Thus, a perceptual change had failed to occur.

3. In terms of their respective variances no significant differences were found between the groups either before or after instruction.

4. No significant differences in variability were found either for the Homeville Group or for the Doe City Group over time.

5. In terms of directionality of variance changes, the Homeville Group indicated a significant trend toward increased variability over time at the .01 level of confidence. Conversely, the Doe City Group indicated a significant trend toward decreased variability over time at the .01 level of confidence.

6. The instructors' perceptions of the stimuli were not significantly different from those held by their respective groups. Neither instructor differed from his group by one or more standard deviations a significant number of times. This was true for both measurements.

7. No differences existed between any of the concepts across scales, and no differences existed between any of the scales across concepts.

Schwartz specified that "to the degree that the sample, the scales, and the concepts are representative,"[1] this research led to the following conclusions:

> First, where the instructor's perceptions of selected course-relevant stimuli are similar to the perceptions of the participants' for the same course relevant stimuli there seems to be little need for instruction. Under such circumstances, one could possibly conclude only that existing perceptions will be re-enforced and that perceptual change will not occur. . . .
> . . . .
> Second, while no perceptual change occurred in the research data, there seems to be some justification for concluding that the group-centered teaching technique reduces the variability of perception to a significant extent. . . . Interpersonal communications in the classroom situation appear to be a necessary requisite to the production of homogeneity in perception.[2]

Of the five foregoing studies conducted under the auspices of University Extension, the research methodology employed in three of the investigatons were both appropriate and sound. The three studies were those reported by Costin and Johnston, by Anikeeff, and by Schwartz.[3] The research methodology used by Murphy may be considered adequate in terms of the data available at time the study was made.[4] The weakness of this study, as is true for many evaluation studies, lies in the fact that evaluation of the program had not been planned at its inception. Thus, initial data were not available on which to base comparisons. Too, a study which uses a mail questionnaire to obtain data is limited in that the investigator must accept the immediate judgments given by respondents. However, Murphy did respect the anonymity of all respondents, and he had assured them of this fact, both of which are considered important factors in obtaining the most valid data from a mail questionnaire study. The research design, per se, which was used by McCormick[1] may be considered appropriate; but since the study was based upon numerical grades earned by students in both groups, it was limited to the differences between the mean grade awarded by the several faculty members.

---

[1] Ibid., p. 55.

[2] Ibid., p. 56.

[3] Supra: Costin and Johnston, p. 146; Anikeeff, p. 148; Schwartz, p. 151.

[4] Supra, p. 139.

The samples used for the five studies may be considered adequate when some limitations are recognized. Johnston and Costin, Anikeeff, McCormick, and Schwartz each used all available participants who were members of the respective adult education courses. Subjects used by Costin and Johnston were those who had been enrolled in the specified course over a period of three consecutive semesters and who could be matched on the basis of identical verbal ability scores. Anikeeff carefully matched both groups of students according to initial performance on each of the tests.

---

[1] Supra, p. 145.

The number of subjects used by McCormick was small as recognized, and he classified students into two groups on basis of those who had completed their extension work before entering the resident program, and those who completed their extension work after they had completed at least one year of their resident work. Thus, groups were not matched. The two groups of students used by Schwartz were approximately matched on three variables. The random sample as originally drawn by Murphy was adequate since it represented approximately 37 per cent of the graduates over time. This is a reasonable density. He matched members of the original sample with members of a control group approximately on three variables in order to compare results obtained from the test which was administered. However, the 85 per cent returns meant that his sample was not completely a randomized sample; hence, one may suspect that his returnees were not a perfect random selection of his population. Murphy was left to his own judgment as to why the other 15 per cent of his sample did not respond, and what difference, if any, their responses would have made on the results.

The quality of the measures used in the five studies varied considerably. The criterion measurement used by Costin and Johnston and by Anikeeff was achievement scores. The instruments which were used to yield these scores may be considered both valid and reliable. Anikeeff reported that his tests had been refined twice by rejection of those items which failed to meet established standards for internal consistency. Schwartz constructed a semantic differential instrument to measure "meaning" and "changes in meaning" which was defined as the criterion measure of teaching effectiveness. This instrument was carefully constructed according to the logic of semantic differentiation and according to the form described by its original authors who had carefully tested their instruments for validity and reliability. Thus, it is reasonable to state that the measures used by Costin and Johnston, by Anikeeff, and by Schwartz were of high quality and were appropriate for stated purposes in each case.

Murphy pointed out that a review of the literature indicated that most firms had developed some technique for evaluating the effectiveness of the various management development processes with which they were satisfied in terms of the validity of the results. Of these techniques, he chose the two which he considered to be most valid; yet, he recognized that the critical incident questionnaire was seriously limited by personal judgments of the rater and by the lack of identical incidents for comparison. The test, "How Supervise?" may be considered a valid instrument. However, Murphy admitted that it had minimum validity for his purposes. Even though he may have used the best instruments which were available to him at the time, it must be recognized that these instruments proved to be of minimum value to him in terms of his stated purposes. Thus, in this case the instruments which were used must be considered of low quality.

McCormick used numerical grade values as the criterion measure. Historically, it is an established fact in education that numerical grade values are likely to be assigned differently by different instructors and that they reflect subjective judgments at least to some unknown extent. Furthermore, the report did not give any criteria upon which the grade values had been based. Hence, the measures used by McCormick are open to serious question and must be considered of poor quality.

The statistical procedures which were employed in the five studies may be considered appropriate. However, Murphy reported results of only one statistical test for significance, but he implied that data had been subjected to tests to determine significance of differences. McCormick did not apply statistical tests of significance to his data. Anikeeff, Costin and Johnston, and Schwartz, each carefully reported results of statistical analysis and used tables to present findings. Anikeeff, in particular, was very meticulous, and the analysis of his data conforms with the highest of statistical standards. In fact, the consistency of results from the two administrations of tests adds more to the confidence that may be placed in the results than he claims. The .05 and higher levels of confidence were used to designate statistical significance of differences in those studies conducted by Costin and Johnston, Anikeeff, and Schwartz. These levels of confidence are generally considered acceptable and are used extensively in research of this type. Murphy reported only the .20 level of confidence obtained which he pointed out indicated a trend, but he made no claim that the difference was a true difference.

The conclusions as drawn in all five studies appear to be sound in terms of stated purposes and of reported findings. In three of the five reports, the investigators qualified their conclusions which indicate that they were cognizant of certain limitations which prevented them from generalizing results to situations other than those studied. The specified qualifications were as follows.

Costin and Johnston made no attempt to generalize beyond the population for their study. Instead, they stated: "The results of the present study have immediate validity only for the particular course, instructor, and populations which were involved."[1] Anikeeff, because of the results obtained in his study, raised a question about the advisability of granting college credit for work performed in evening off-campus extension courses. He also stated: "If further studies support results found in this investigation, the case for granting college credit for extension work will be severely challenged."[2] Schwartz, in drawing his conclusions, prefaced them with the statement: "To the degree that the sample, the scales, and the concepts are representative, the results of this research lead to the following conclusions."[1] Furthermore, he pointed out that "the research design used limits the degree of generalization that can be made. Restricted field studies are notorious for the limits imposed on generalizations."[2]

---

[1] Costin and Johnston, op. cit., p. 126.

[2] Anikeeff, op. cit., p. 173.

Four of the five studies may be considered to contribute to practice. The four studies were conducted by McCormick, Murphy, Costin and Johnston, and Anikeeff. One of the five studies, conducted by Schwartz, definitely contributes to knowledge.

The study by McCormick was based upon numerical grades earned by the two groups of students; hence, it is limited to the differences between the mean grades awarded by the several faculty members. The results report the findings for this one specified comparison and offer no basis for generalization to other situations or even to future groups within this university. The value of this study will be dependent upon its contribution to instructional and grading practices at the University of Virginia, plus whatever stimulation it may offer others engaged in similar programs to evaluate their own practices. Neither the results, nor the techniques, make significant contributions in the field of extension instruction or research methodology. The study by Murphy contributes to practice. Its value lies in the fact that the data obtained from the forty-six supervisors before training may be used for comparisons of results after training.

---

[1] Schwartz, op. cit., p. 55. [2] Ibid., pp. 57-58.

The two studies by Costin and Johnston and by Anikeeff, respectively, each contribute to immediate practice in situations that may be judged to be highly similar to that in which their studies were conducted. The limitations of their samples so restrict the general applicability of their findings that their contribution to general knowledge is not significant in, and of itself. However, each of these two investigations is a carefully controlled study of a single behavioral outcome of a specific instructional technique in a specific subject field for two sets of students. Each study, especially the one by Anikeeff, is a model that deserves to be emulated. Results from repeated studies conducted by other instructors with other students could be consolidated to obtain a definitive evaluation of the instructional technique (lecture and text) when applied generally in the respective subject areas or in other content areas. Student characteristics should be reported as carefully as in the Anikeeff study.

The study by Schwartz contributes to knowledge in that he applied a new theory of evaluation to a specified adult education program. The development of a criterion measurement based on the logic of semantic differentials opens up a theory which may be tested in other situations.

Evening Colleges

Two studies were identified which relate to the occupational function of Adult Education and which had been conducted under the auspices of Evening Colleges. Each of the studies evaluated the evening college program by comparing selected aspects of the program with that of the regular day program. Thus, the basic research designs were similar. However, one of the studies represents a higher level of research because of a pilot study conducted for the purpose of establishing the validity and reliability of the instruments which were used to obtain data. Because of this factor, and since only two studies were identified, both studies are reported. An analysis of each study is made, after which a summary evaluation of the two studies is presented.

Gibson compared regular day students with continuing education students in terms of selected factors and of attitudes.[1] He stated the following hypotheses:

1. There are no significant differences in course achievement, as measured by grades, of students who are enrolled in day courses when compared with students enrolled in continuing education classes.
2. The number of academic units a student is enrolled for in continuing education courses has no adverse effect on the course grade.
3. A student's age and maturity level has no effect on course grade in continuing education courses.
4. Employed students attain equal grades when compared with the non-employed student.
5. Length of class meeting does not affect students' course grades.
6. There are no differences in the amount of time per week continuing education students and day students spend on assignments.[2]

Before proceeding with the primary study Gibson conducted a pilot study to test the validity and reliability of the instruments which were to be used to obtain comparative data. The instruments were identified as the "Students' Questionnaire" and the "Faculty's Opinionnaire." The report contained a lengthy discussion of the procedures involved in: (1) the development of the instruments, (2) the conduct of the pilot study, and (3) the establishment of the validity and reliability of the questionnaires. These procedures are reported briefly in the subsequent paragraphs.

---

[1] James Edward Gibson, "A Study of Selected Factors and Attitudes Comparing Regular Day Students with Continuing Education Students" (unpublished Ed.D. dissertation, University of Arizona, 1962).

[2] *Ibid*., pp. 13-14.

The students' questionnaire elicited both biographical and attitudinal information. It was divided into two sections: A and B, each of which yielded similar data. Section A, intended for both day and continuing education students, contained thirteen statements designed to educe students' attitude toward course difficulty. Section B, intended only for continuing education students, contained a series of statements designed to indicate general attitudes toward continuing education offerings. The faculty opinionnaire consisted of thirty-five statements formulated to elicit faculty attitudes toward day and continuing education classes. Gibson reported his rationale for the inclusion of each item in both instruments. Responses to statements in the instruments were made according to scale devised by Likert: "strongly agree," "agree," "no opinion," "disagree," "strongly disagree."

The pilot study was conducted during the second semester of 1960-61. The student questionnaire was administered to 101 students who were enrolled in six different classes. The faculty opinionnaire was administered to fifty faculty members who either were teaching or had taught late afternoon and evening classes during the previous semester.

Since most of the biographical type of information could be checked for its authenticity in the Registrar's office, only the phrasing of the items was examined after this administration. Based on suggestions received from the 101 students and from professors in the College of Education, minor revisions were made. An item analysis was made of the responses to all statements in both instruments. The Phi Coefficient was computed for each scale of each statement and results were presented in tables. Of the thirty items in the student questionnaire, all except six were found to be discriminating at the .05 level of confidence. For the faculty opinionnaire, all items were discriminating at the .05 level of confidence. The reliability of the instruments was determined by the test-retest method. The student questionnaire was re-administered to sixty-seven students enrolled in three of the same classes one week following the first administration. The faculty opinionnaire was re-administered to twenty-five of the fifty instructors within a period of ten days to two weeks following the first administration. Results from the computation of the Pearsonian r for each instrument yielded a reliability of .91 for the student questionnaire and .86 for the faculty opinionnaire.

Participants for the primary study were selected on the bases of certain criteria. First, only those courses which were taught by the same professor both before 4:30 p.m. and after 4:40 p.m. were considered. The catalogue listed ninety-two such classes but seven did not materialize because of an insufficient number of students. Secondly, only those courses which were comparable in terms of content material, techniques of class presentation, and evaluative procedures were included. The investigator interviewed professors to determine if classes met this criterion. Six courses were disqualified because professors indicated dissimilarity in course content. Finally, all professors whose courses met the two criteria were contacted; they were briefed on the purpose of the study and requested to participate. All but two faculty members agreed to participate. Faculty members who agreed to participate also consented to: (1) provide time for two personal interviews with the investigator, (2) make examination scores and other grades available to the investigator, and (3) relinquish twenty minutes of class time for the administration of questionnaire to students enrolled in both the day and evening sections. Accordingly, the subjects of this study included those students who were enrolled in the courses so selected and the corresponding professors. Instruments were administered to participants as indicated which included 2,035 students. Final student comparisons were based on sixty-nine courses and thirty-two professors. However, a total of fifty professors were interviewed, and since all items in the faculty instrument were found to be discriminating, responses to both the interview and opinionnaire were included.

The collected raw data were coded for IBM tabulations. To compare students in day classes with those enrolled in continuing education classes, students were divided into two categories: (1) those who were enrolled in classes which met during the period of time from 7:00 a.m. to 4:30 p.m., and (2) those enrolled in classes which met during the time from 4:40 p.m. to 10:30 p.m. Day classes met three fifty-minute periods a week and evening and night classes met fewer times a week with a prolonged period. Of the 2,035 students, 1,623 were classified as regular day students and 414 as continuing education students.

Some of the data were statistically analyzed. For each group, mean scores, standard deviations, and standard error of the means were computed. The t-test was employed to determine if differences in scores for the two groups of students were statistically significant. The .05 and higher levels of confidence were designated for establishing significance. Other data were reported in percentages.

The report contained a lengthy descriptive interpretation of the findings. Many tables, presenting data and the results of analyses which were made, were appropriately arranged within the context. The investigator reported in descriptive fashion several characteristics of both students and faculty members. For the students, reported characteristics included sex, marital status, and class, enrollment, degree, employment, and age classifications. For faculty members, reported characteristics were age, number of years in teaching, number of years teaching at the University, earned degrees, and academic rank. It was determined that the study included all classifications of students; however, proportionately more juniors were involved. Four colleges and fourteen different departments of the University were represented. The number of employed students was 760, of which 483 were day students and 277 were continuing education students. The mean cumulative grade point average of students in the study was 2.645. This average was .23 of a point above the mean cumulative grade point average of the total university population.

Of the many findings which were reported, the following appear to be of most importance since they relate specifically to the stated hypotheses:

1. Continuing education students made significantly better grades than did the day students (.05 level of confidence).

2. No significant differences existed in mean average grades between day and continuing education courses and units carried with the exception of students who carried ten to twelve units of credit. Gibson specified that full-time employed continuing education students should not be permitted to carry more than seven units since their grades were affected negatively after this number of hours.

3. Null hypothesis number three was accepted only for the forty-one to fifty age group (.05 level of confidence). Continuing education students under twenty years of age had a significantly lower grade point average than did corresponding day students. Enrollees over twenty-one years of age, with the exception of the thirty-one to forty-one age group, performed significantly better in continuing education classes than did the corresponding day enrollees.

4. Employed students made significantly better grades than did the nonemployed students (.05 level of confidence). No significant differences were found in the grades received by day students who were employed and continuing education students who were employed.

5. Neither the time of day nor the prolonged periods of time for classes had adverse effect on students' grades.

6. Continuing education students studied significantly more hours than did day students.

Gibson pointed out that the attitudinal section of the study raised more questions than it answered; therefore, it was somewhat disappointing. Concluding his summary, he stated: "The data presented here do give some basic research in the areas of student attitudes, study habits, course achievement, and professional attitudes toward continuing education."[1]

Five recommendations were made to improve the continuing education program. Gibson based his recommendations upon the evidence that: (1) continuing education students received significantly higher grades than did the day students, (2) employment appeared to enhance class achievement, (3) attitudinal data indicated students had no difficulty in studying for continuing education courses of prolonged class periods, and (4) continuing education students studied significantly more hours than did day students.

Manley identified certain characteristics of the extended day program at San Fernando Valley State College and ascertained the quality of its program as compared with the regular day program at the same institution.[2]

---

[1] Ibid., p. 91.

[2] Charles Arnold Manley, "An Evaluation of the

Comparisons of the extended day program with the regular day program were made in terms of: (1) faculty training and experience, (2) grading standards and practices, (3) faculty opinion, and (4) student opinion.

Data were obtained from two sources: (1) official records of the college which included admission and personnel records, publications, and minutes of college committee and faculty meetings, and (2) three specially constructed survey instruments identified as, "The Extended Day Student Questionnaire," "The Extended Day Student Opinionnaire," and "The Questionnaire for Instructors Teaching Evening and Day Sections of Same Course." These instruments were used as follows. The Extended Day Student Questionnaire, designed to obtain basic census-type information about students enrolled in the extended day program, was administered by instructors to students in attendance during the week of January 11-16, 1960. It was completed by 2,163 of the 2,637 students who were enrolled in extended day classes during the fall semester 1959-60. The Extended Day Student Opinionnaire, designed to elicit opinion regarding the two programs, was mailed to the 513 students who had been

---

Extended Day Program at San Fernando Valley State College" (unpublished Ed.D. dissertation, Washington State University, 1961).

enrolled in both programs. Returns were received for 429 students. Faculty opinion was obtained by the questionnaire from 24 of 25 instructors who taught regular day and extended day sections of identical courses concurrently during the specified semester. This questionnaire was designed to yield information on the extent to which instructors felt that the two programs paralleled one another with emphasis on such matters as course structure, estimated calibre of students, teaching procedures, and the extent of student learning.

The data obtained from the three survey instruments were coded for IBM tabulations. Characteristics of the extended day students which were identified included, among others, sex, age, marital status, outside employment, college background, matriculation status, class load, class attendance, length of enrollment, and educational objectives. Such characteristics as educational background, academic rank, and tenure were identified for the faculty members.

Some of the data were analyzed statistically. The chi-square test was applied to distributions to determine significance of relationships. The .05 and higher levels of confidence designated significance. Other data were reported in numbers and in percentages. The report contained a lengthy description of the history and characteristics of the extended day program and of the comparative findings. Thirty-three tables which presented the analyses of data were interspersed within the context. The following is a brief summary of major findings:

1. No significant difference in academic training, as measured by highest earned degree, was found in the regular day and extended day distribution of instructional units among the three academic degree classifications. However, significantly fewer instructors of higher academic rank and longer tenure were teaching extended day classes.
2. No significant differences in "A" through "F" grade distributions were found in six of the seven academic divisions. Nonetheless, significantly more students withdraw from extended day classes (.001 level of confidence).
3. The vast majority of the faculty who had taught concurrent sections of the same course in both programs indicated use of similar teaching procedures and evaluation instruments, and the maintenance of comparable academic standards in both programs. There were twice as many instructors who reported student achievement higher in the regular day sections as there were who found the same true in extended day sections. Two-thirds of instructors indicated more advisement and counseling time was spent with regular day students. Over one-half found better attendance in the day classes.
4. A majority of the students reported no difference in the two programs in terms of subject knowledge of instructors and of their abilities to present the material effectively.

Manley concluded that, in general, the extended day program at San Fernando Valley State College was comparable to the regular day program. However, because of certain trends in the data, he made the following recommendations:

1. Schedule as many extended day courses as possible to meet twice weekly.
2. Increase availability of competent advisors during extended day hours.
3. Distribute more equitably instructors of higher rank and more experience to the extended day program.
4. Extend library hours to better accommodate student personnel services during extended day hours.[1]

The foregoing two studies conducted by Gibson and by Manley evaluated evening college programs by comparing selected aspects of the programs with those of the regular day programs. The research methodology employed by both investigators may be considered appropriate in terms of stated purposes and in terms of data which were available at the time.

Gibson's study represents a higher level of research than did the one conducted by Manley. In fact, Gibson conducted two studies. The processes involved in the construction of an instrument, in the development of a theoretic rationale for including each item, in the conduct of a pilot study for the purpose of testing the instrument for validity and reliability, and in the application of statistical analyses to determine the extent to which items possess discriminating power and to which the instrument is reliable is an acceptable piece of research within itself. However, Gibson chose to use the instruments to compare the two groups of students in terms of the selected factors and attitudes which the instruments were designed to measure.

---

[1] _Ibid_., p. 117.

In both studies samples were more than adequate in terms of numbers. However, neither investigator made any effort to match the two groups of students. Instead, they simply used all available subjects who had been enrolled in both the regular and extended day programs (Manley) or who were enrolled in comparable courses which were taught by the same professor (Gibson). Much data on the characteristics of subjects were obtained by both investigators, but they were not used for matching purposes. It is the judgment of the present investigator that the quality of the research would have been enhanced had the researchers used smaller but carefully matched samples of participants.

The quality of the measures varied for the two studies. Admittedly, the questionnaires used by Manley were designed to elicit basic census-type information and to obtain opinions in relation to the two programs. He reported that statements included in the instruments had been revised after evaluation of the instruments by appropriate professional people, but he did not indicate that any effort had been made to establish validity and reliability. Manley did obtain other data from official records, but it is reasonable to assume that they were used primarily to report the history and the characteristics of the extended day program. However, some of the data may have been reinforcing to the census-type information obtained from one of the questionnaires. Nonetheless, it must be recognized that the measures used by Manley were of poor type and quality.

Gibson made a careful and conscientious effort to develop valid and reliable instruments for his purposes. Because of the high discriminating power of the items and the approximately .90 reliability coefficient which were established, one is justified in stating that the type and quality of measures used by Gibson were of high quality for his population. However, it must be pointed out that these measures have validity only for the specified population of the one university. Hence, other researchers must first test these instruments with a sample of their population in order to establish their usefulness in a different situation.

Some of the data were analyzed statistically in each of the two studies. Gibson employed the t-test to determine significance of differences between mean scores and Manley applied the chi-square test to determine significance of relationships. Confidence levels were established at the .05 and higher levels by both investigators.

Conclusions for the two studies appear to be sound in terms of stated purposes and of reported findings. Each investigator made specific recommendations for immediate improvement of the specified program. However, the results of the studies have immediate validity only for the particular courses, professors, and populations which were involved.

The two studies contribute to immediate practice for the two institutions. Since each of the studies was limited to a specified institution for a specified time, the results are not generalizable to other institutions nor to the same institutions over time. The value of the studies lies in the fact that data may serve as the basis for future comparisons.

Junior Colleges

Only one study was identified which relates to the occupational function of Adult Education and which had been conducted under the auspices of a Junior College. This study made no pretense of evaluating a program, per se; rather its purpose was to analyze critically a total adult education program in a Junior College. Hence, one may question the inclusion of this study in this report. However, since it was the only study identified, and since it represents a serious effort to formulate criteria for analyzing a junior college adult education program, the study is included. It may contribute to further, and more extended work in evaluation of Junior College Adult Education Programs. The analysis and evaluation of this study is reported in the subsequent section.

Luchsinger made a critical analysis of the Adult Education Program of Kilgore College which was based on criteria formulated from data submitted by 124 Junior Colleges in the United States.[1] The four purposes of his study may be stated as follows:

1. To determine the gradual evolution of adult education as a function of the junior college from a review of the literature of the history and development of junior colleges and of the adult education movement;
2. To compile and analyze data obtained from junior colleges in terms of: (a) adult enrollment, (b) adult student personnel services, (c) administration of the adult program, (d) the position of the adult education administrator, (e) instructional practices

---

[1] Leland Beyer Luchsinger, "Criteria for Evaluation of Adult Education Programs in Junior Colleges and Their Application to Kilgore College" (unpublished Ed.D. dissertation, University of Texas, 1959).

and procedures, and (f) curricula in junior colleges;

3. To develop criteria for analyzing an adult education program in a junior college; and
4. To apply criteria to the adult education program of Kilgore College.

Luchsinger selected from the 1957 Junior College Directory those junior colleges which had at least 100 adult and/or special students enrolled. A return postal card was mailed to the 248 junior colleges so selected. Responses to two questions were solicited: (1) "Does your college have an evening division or an adult education program?" (2) "Will you please complete a questionnaire pertaining to adult education if and when one is sent to you?"[1] Of the 183 replies received, 151 responded affirmatively, and 32 gave negative replies or stated that a questionnaire was not applicable to their program. An inquiry form, "A Questionnaire for Establishing Criteria to be Used in Analyzing an Adult Education Program in a Junior College," was mailed to the 151 colleges which had responded affirmatively. The questionnaire elicited detailed information pertaining to such matters as philosophy, goals, enrollment, student personnel services, organization, administration, instruction, and curriculum of the adult education program. Replies were received from 127 junior colleges, but three questionnaires were not usable. Thus, the information obtained from 124 junior colleges was used as the basis for the study.

---

[1] Ibid., Appendix A.

Data were tabulated and reported in sixty-one tables and in twenty-nine figures. The range, median, and quartile point for the data were presented in the majority of tables. The tabulated responses were classified in six major divisions: (1) enrollment, (2) student personnel services, (3) administration of the adult education program, (4) position of the administrator of adult education program and his staff, (5) instructional practices and procedures, and (6) curriculum. Each of these major divisions contained several sub-divisions, each of which was carefully analyzed in order to establish the criteria for the stated purpose. The magnitude of this task may be illustrated by summarizing briefly the manner in which two divisions were analyzed.

The enrollment data were tabulated in three categories: (1) 0 - 499, (2) 500 - 999, (3) 1000 and over. These tabulations were analyzed in terms of: (1) ratio of adults enrolled to each day student, (2) total number of adults enrolled, and (3) ratio of adult students or adult population to total population of the college community, and

(4) the percentage of adults attending adult classes. Seven factors were considered in the analysis of data obtained on student personnel services. These were: (1) student organizations, (2) services of student center or student union, (3) services rendered by the college book store, (4) the counseling and testing services rendered adults by the guidance program, (5) the administrative procedure, (6) the service of the college library, and (7) the follow-up procedures employed to assist adults.

Accordingly, all data obtained from the 124 colleges were carefully analyzed. A description of the criteria used in the analysis of these data and the interpretations of the derived tables and figures were presented in chapter three of the report. This chapter, which also included the sixty-one tables and the twenty-nine figures appropriately placed within the context, consisted of 164 pages.

Criteria for analyzing an adult education program in a junior college were deduced from the data in the tables and figures. A lengthy, descriptive criteria-analysis instrument was formulated. The instrument was organized in six major divisions as follows:

1. Criteria for analyzing the philosophy and objectives of the adult education program;

2. Criteria for analyzing the size of the adult education program;
3. Criteria for analyzing the services rendered adult students;
4. Criteria for analyzing the organization, administration, publicizing, and financing of the adult education program;
5. Criteria for analyzing the adult administrative position and his staff; and
6. Criteria for analyzing the instructional procedures, practices, and curriculum.

The several criteria listed under each division were based on accepted procedures employed by the surveyed junior colleges. Five descriptive statements indicated the bases for rating each criterion as "unsatisfactory," "poor," "fair," "good," and "excellent," respectively.

The bases for the ratings were derived from the majority of the tables as follows. The lowest point on the range was used to designate the "unsatisfactory" rating, and the highest point on the range was applied for the "excellent" rating. The first quartile yielded the "poor" rating; the third quartile served for the "good" rating, and the median was employed for the rating of "fair." The bases for ratings derived from the other tables which revealed only the frequency of mention were as follows. Statements composing a table were listed sequentially from the one which was mentioned the greatest number of times to the statement mentioned the least number of times. An "unsatisfactory" rating was given to a criterion when a minimum number of statements were checked, and the maximum number of checked statements represented the "excellent" rating. The rating of "good" was assigned when three-fourths of the statements which comprised a table were checked; the rating of "fair" resulted when one-half of the statements were checked, and the rating of "poor" was specified when one-fourth of the statements were checked.

The second division of the instrument, "Criteria for Analyzing the Size of the Adult Education Program," is described briefly in order to illustrate items and the format which were used. This division contained eight different criteria, the first of which was "Ratio of the number of adults enrolled to each day student." The descriptive statements on which the ratings were based for this criterion represent the briefest kind of statements which were included in the instrument. They were as follows:

__unsatisfactory	.25 adults to one day student
__poor	.75 adults to one day student
__fair	two adults to one day student
__good	3.75 adults to one day student
__excellent	seven adults to one day student[1]

---

[1] *Ibid*., direct from the instrument, pp. 324-332.

The number of items which were included for each division of the instrument were as follows. For division I, there were three; II - eight; III - ten; IV - eleven; V - three; and for VI, there were ten. The final instrument consisted of twenty-seven pages, a copy of which was included in the report.

Luchsinger pointed out that "these criteria are to be used in analyzing junior college adult education programs; they are not evaluative criteria."[1] He continued with his explanation:

> An analysis of an adult education program necessitates differentiating between the composites, discovering their relationship to one another, learning which practices, functions, procedures, methods, techniques, records, reports, facilities, and services are necessary and which can be eliminated, revised, or maintained. These analyses are based on standard practices used by the 124 responding junior colleges. No attempt was made to compare one program with another.[2]

The instrument was administered to nineteen people who rated the adult education program of Kilgore College. These individuals were the president, the dean, the adult director, and sixteen adult education instructors. Also, seventy-eight adult education students rated the program in terms of twenty statements which were selected from the instrument because they pertained to adult students. Of the total of 196 usable ratings, four were "unsatisfactory," ten were "poor," sixty-four were "fair," sixty-seven were "good," and fifty-one were "excellent." Luchsinger concluded that from this analysis the adult education program of Kilgore College could be compared, in most instances, with the median or better adult education programs as offered by a sampling of 124 adult education programs in junior colleges. This conclusion follows from the fact that 182 of the 196 ratings fell at, or above "fair," the median of the 124 college programs studied.

---

[1] Ibid., p. 243. [2] Ibid.

Essentially, the study by Luchsinger was designed to develop an instrument. This is a logical first step in research in this area. The investigator made a careful and conscientious effort to develop such an instrument. He developed from the reported practices in a sample of 124 junior colleges a series of scales in terms of which other adult education programs in junior colleges could be analyzed.

In the process of developing this instrument, Luchsinger analyzed meticulously the vast amount of data which were obtained from the responding junior colleges. These data were categorized and reported appropriately and adequately in tables and figures from which the ranges, medians, and quartile points were computed. He gave special attention to the quality of data he used to develop the series of scales which made up the final instrument. In defining the positions on each scale, he was able to develop descriptive statements which identified rather precisely each of the five positions, namely, unsatisfactory, poor, fair, good, and excellent. Thus, the adult education program could be analyzed in terms of its rank based upon the composite of the practices employed by the 124 junior colleges.

This study should make a very definite contribution to the analyses of practices in junior college adult education programs. More importantly, the product of this research should be useful in further research designed to evaluate adult education programs sponsored by junior or community colleges.

Public School Adult Education

Of the evaluation research conducted under the auspices of Public Schools, two studies were identified which relate to the occupational function of adult education. Both studies are descriptive in nature. Since they differ in terms of methods used, both are reported. Each of the studies is analyzed, after which a summary evaluation of the two studies is made.

Cheskie conducted a study to appraise the current adult education program in the Highland Park Public School System and to make recommendations for further program development.[1] The program was appaised in terms of: (1) student evaluation of the general program, (2) student interpretation of needed services, and (3) community leader appraisal of the general program and of needed services. Recommendations for further development were made in terms of: (1) guidance and counseling, (2) curriculum, (3) instruction, and (4) special services.

Primary data were obtained from a questionnaire administered to a sample of participants. To ascertain community reaction, additional data were obtained during personal interviews with twenty-two community leaders who were selected on the basis of their active leadership in the community. The questionnaire, specially constructed and pretested, elicited personal information, reactions to the program in general, reasons for entering the program, and recommendations for program development. It was administered by teachers of adult education classes to 2,096 participants within a one-week period during the winter term, 1957. Of the 1,892 questionnaires returned in usable form, a random sample of every other one was drawn. Thus, the final sample included 946 adults who were enrolled in classes during the specified term. The investigator conducted interviews with twenty-two community leaders who were randomly selected from a list of forty-four leaders representing the various influential groups in the community.

---

[1] Sophie Veronica Cheskie, "An Appraisal of the Adult Education Program of the Highland Park Public Schools, Highland Park, Michigan" (unpublished Ed.D. dissertation, Wayne State University, 1957).

Responses on questionnaires were coded and categorized so that certain relationships could be studied and appraised. A great number of tables presented the data in frequencies and in percentages. Relationships were established by inspection. The lengthy descriptive analyses, supported by the tables, related the personal characteristics of the respondents to certain aspects of the program. Personal information was classified in six major categories: (1) educational background, (2) employment and occupational status, (3) income, (4) marital and family status, (5) age and sex, and (6) residence and home ownership. Sequentially, each of these six categories was related to eleven aspects of the adult education program. The eleven aspects were:

1. The guidance and counseling services which were found to be of most value;
2. The reasons given for enrolling in adult education classes;
3. The reasons given why some people did not enroll in adult education classes;
4. The opinions expressed on desirable personality characteristics of adult education teachers;
5. The opinions expressed on desirable teaching techniques;
6. The opinions given on undesirable personality characteristics of teachers;
7. The opinions given as to undesirable teaching techniques;
8. The kinds and number of classes in which participants currently were enrolled;
9. The kinds and numbers of classes which had been taken previously;
10. The kinds and numbers of classes which were to be taken in the future; and
11. Additional suggestions in terms of classes and activities which should be incorporated in the program.

Thus, sixty-six relationships were analyzed and reported. These analyses constituted the greater portion of the 517-paged report. Moreover, in addition to nineteen tables which were presented in the context, 245 pages of the Appendices contained additional tables which supported the analyses. Of the many findings which were reported, the following are illustrative;

1. In terms of curriculum offerings, she found that four course areas had been most popular. These were: "vocational," "educational," "social," and "leisure-time" activities. The heaviest enrollments had been in "general academic" classes and in "leisure-time" activities. The greatest future demands were for the "general academic" and "business and commercial" courses.

2. The community leaders thought of the public adult education program as generally meeting the present needs of the community; they considered instruction adequate and the administrative leadership superior. For future development, they suggested a continuation and expansion of the basic program within the current curriculum framework.

3. She found that persons employed in the professional, craftsman, and laborer classifications were more aware of definite objectives and had concrete reasons for enrolling or for not enrolling in adult education classes, than were those in other occupational classifications.

Cheskie drew eleven conclusions which were based on the analysis of the returns of the students' survey. The following were included:

2. The most popular reasons for enrolling in adult educational classes were "vocational," "cultural," "educational," and "social."

. . . .

6. Students were currently enrolled in the following subject areas: "general academic," "business and commercial," "homemaking," and "leisure-time" activities.

. . . .

10. There was no increased demand, on the part of current enrollees, for classes in "civic and public affairs," "personal improvement," "parent and family life education," and "industrial training." However,

> there was mention of some unsolved problems or "worries" involving personal and financial difficulties for which there was hope to find assistance through adult education.[1]

Six conclusions were based on the analysis of the interviews with community leaders. The following were included:

> 1. The adult education program was evaluated as meeting adequately the present needs of the community, but must expand the present course offerings into more sequential levels of work.
>
> . . . .
>
> 5. The curricular offerings were conceived of as having wide variety of informal classes but as lacking in diversification in the formal and academic areas.[2]

Cheskie made seventeen recommendations, organized under the four areas as proposed, for the improvement of the adult education program in the Highland Park Public School System. Included were the following:

> 1. There should be a continued strong program in the "general academic" and "business and commercial" subject areas.
> 2. There should be an increase in co-sponsored classes with business, industry, government, and other community groups. . . .
> 3. There should be continued efforts to articulate the philosophy and importance of adult education not only to faculty but to all leaders in the community.[3]

Timken conducted an evaluation study which was similar to the previous investigation in that it was descriptive in nature.[1] However, the purposes of the study as stated in his report were inconsistent. He stated in the introduction that:

---

[1] Ibid., pp. 236-238.

[2] Ibid., p. 238.

[3] Ibid., pp. 241-242.

> This study is an attempt to establish a process by which public school adult education programs may be studied. The study does not present an attempt to reach a final set of program objectives, methodology, or organizational procedures, but the study has gathered from literature and practices found in the field a pattern upon which a process to study adult education may be built.[2]

Then, in Chapter I, he stated:

> The purpose of the study is to determine the adequacy of three selected Oklahoma public school adult education programs in relation to proposed goals and the expressed needs of their students. In addition, the study proposes to accomplish the following secondary purposes:
> 1. To propose goals for public school adult education programs which emphasize the educational processes conducive to effecting desirable changes in people.
> 2. To list specific adult education objectives which can be used for forming the educational activities for public school adult education programs.
> 3. To determine specific educational areas of problem concern held by the adult students enrolled at the time of the study in the three selected public school adult education programs.
> 4. To survey the three selected public school adult education programs in regard to (a) educational

---

[1] Joseph E. Timken, "A Study to Determine the Adequacy of Three Selected Oklahoma Public School Adult Education Programs" (unpublished Ed.D. dissertation, Oklahoma Agricultural and Mechanical College, 1951).

[2] *Ibid*., p. viii.

methods and organization of programs which when used tend to meet the proposed goals, and (b) emphasis being placed upon various specific adult educational objectives listed in the adult problem questionnaire.

5. To state conclusions and make suggestions for the improvement of the public school adult education programs being studied.[1]

Each of the three public school adult education programs selected for evaluation had a decided difference in emphasis. The Junior College at Miami, Oklahoma, was a typical public school program in that it operated academic and trade and industrial education programs through the thirteenth and fourteenth year. The Stillwater, Oklahoma, Public School Adult Education Program placed emphasis on the general approach toward all interested members of a single public school community. The Tulsa Public School Adult Education Program was that of the typical night school as conducted by public schools in such other cities as Denver, Baltimore, Milwaukee, and Cleveland.

Following the identification of the three selected public school programs, it was pointed out that the specific needs to be met by this study were assistance with the solution of those problems which had kept public school administrators from launching an active, coordinated plan of adult education. These were:

---

[1] *Ibid.*, p. 1.

1. A guiding philosophy based upon open-end goals which can bring understanding and direction to teachers and adults involved in forming and conducting programs of public school adult education.
2. An instrument with which the interests of adults can be determined. This is an individual approach to personal problems. The recording and activating of such expressed needs are probably the most important administration problems to be met.
3. The understanding of the problems of the adult in relationship to his being a learner. These problems of adults should be accommodated by organizational adjustment. These are known but seldom considered as guides to the arrangements for learning activities for adults. There is basic research available which points to the necessity of a variety of approaches and methods in a successful adult education program.[1]

During the conduct of the study Timken approximated a pattern of investigation which may be summarized as follows:

1 He reviewed the literature, and he selected the proposed goals after a careful study of the most critical literature.

2. He interviewed a large number of adults to check the areas of problem concern so that the instrument which was developed and used (a check sheet) would have the common problems expressed in meaningful terms to the respondents.

3. He made additions to a plan of public school adult education program after submitting it to Dr. Homer Kempfer of the United States Office of Education, who compared items on study sheet with a similar study in progress in terms of the approaches and methods listed to meet the proposed adult education goals.

---

[1] *Ibid.*, p. 4.

4. He visited public school adult educators to study cooperatively the proposed goals in order that the study questionnaire would present a true picture of the attempts of the selected programs to reach these goals.

5. He compiled a list of adult problems as a result of reviewing other studies and of interviewing adults who made lists of problems that they had experienced. This list was submitted to the adults enrolled at that time in the public school adult education programs under consideration. Each respondent indicated the extent of his concern about each problem by checking the scale as follows: 0--no concern, 1--some concern, and 2--definite concern.

6. The list of problems was checked cooperatively by the directors and teachers of adult education programs under study. Respondents indicated the extent of emphasis being given to the problems by checking the scale as follows: 0--no emphasis, 1--some emphasis, and 2--a great deal of emphasis.

Timken proposed five goals for public school adult education. These goals were conceived in terms of desirable changes in people that should result from the program. They were: (1) to create a desire for equality of educational opportunity, (2) to develop a self-directing learner, (3) to develop an informed, constructive thinker, (4) to encourage the adult to become a cooperative being, and (5) to guide the adult into experiencing a continued growth in adjusting to his environment and to the people of his community and nation.

He selected educational objectives which were based upon the specific problems of living. These objectives, according to the investigator, would enable the adult to adjust to his environment and to the people of his community and nation. They were as follows:

1. Vocational Competency: To enable all individuals to gain the status of a producer of desired goods.
2. Family Living: To develop desirable relationship of members of the family in consumer, moral, social, and civic activities.
3. Maintaining Health and Safety: To acquire proper attitudes, habits, knowledge, and skills in the fields of nutrition, exercise, mental attitudes, care of body, disease, and dangerous situations found in life.
4. Increasing the Fundamentals of Learning: To increase the ability to read with more understanding and to communicate more effectively by speaking and writing. The increase of the knowledge and the use of needed mathematics. The improved use of the problem-solving technique.
5. Local, State, National and International Citizenship: To develop the understanding and the proper participation in affairs that relate the individual and his welfare to the movements of government.
6. Expression by Creating or Appreciating Beauty: To gain skill and understanding in activities that give pleasure and allow for the release of nervous tension; also, the development of the individual's contribution to win group's acceptance.
7. Adjusting to the Several Life Stages: To enable individuals to understand and gracefully adjust to the changing responsibilities, relationships, and physical and mental demands of the succeeding life stages.
8. Spiritual Balance: To enable people to attain peace within themselves by understanding and adopting a religious creed to live by--an alliance with a group that holds the same religious creed. The ability to pass on this inner light to their children or to their associates.[1]

---

[1] *Ibid*., p. 27.

Subsequently, he formulated criteria which were based upon accepted educational procedure. These criteria indicated the necessary steps to be taken in order to achieve the public school adult education goals as proposed. He proposed to use them as a basis to evaluate the three selected public school adult education programs. The criteria were listed as sub-heads to the goals in order to give continuity to the measuring instrument as follows:

Goal I. To Create a Desire for Equality of Educational Opportunity

a. All educational leadership involved in the public school adult education program should be active in the identification of needs and interests of the community's adults.

b. Needs and interests should be determined in such way that the process points out the importance, to the adult, of his becoming a learner and implements the organization of such learning activities for adults.

c. The total of all means of communication should be used to contact the adults of the community.

d. That a plan be used, in the determination of needs and interests, that includes all segments of a community's adults, and their needs and interests.

Goal II. To Develop a Self-directing Learner

The process used in the learning situations for adults should include teacher-pupil planning, the use of the problem-solving process, the evaluation of results, and the producing of a state of mind on the part of the adult student to continue on to further learning.

Goal III. To Develop an Informed, Constructive Thinker

The adult should have proper information available through all avenues of formal and informal education.

The agencies and institutions of the

community that are active in movements toward better living should give direction, through all media of communication, as to the facts which should guide the decisions of its citizens.

Goal IV. To Encourage the Adult to Become a Cooperative Being

- a. That there be cooperation of all those concerned in the educational processes within the public school adult education program.
- b. That there be cooperation between the public school adult education program and all agencies and institutions of the community which have adult education or public relations needs.

Goal V. To Guide the Adult into Experiencing a Desirable, Continued Growth in Adjusting to His Environment, and to the People of His Community, Nation, and World in a Democratic Manner

- a. The felt needs of the adults of the community should be determined.
- b. The needs once determined should be recorded.
- c. The needs should be listed according to person and together with appropriate needed learning activities.
- d. Learning activities should be organized when enough adults are recorded.
- e. The learning experiences should be properly guided as to organization, methods, and direction. The evaluation process should be used for encouraging future growth of the adult.
- f. The adults' records should constitute the information necessary for a guidance folder.[1]

Timken constructed two measuring instruments to use in evaluating the three selected public school adult education programs: (1) The Study Guide, and (2) The Check List. The Study Guide, based on the criteria listed above, was used to determine the extent to which practices were employed in the given programs. Organized in seven sections, this instrument included from six to thirty-one appropriate practices under each of the five goals and included additional questions which elicited institutional and historical types of data. The director in cooperation with teachers in the several departments responded by checking a similar scale for each section: 0--seldom used, 1--sometimes used, and 2--often used.

---

[1] *Ibid.*, pp. 28-29.

The purpose of the check list, a questionnaire, was to determine the problems of adults for which there was a felt need for more knowledge, skill, or understanding, It was carefully constructed from the list of specific objectives and contained more than ninety items. Enrollees rated each item using the following scale: 0--if no problem, 1--if some concern felt, and 2--if a definite problem. The check list also was submitted to the directors of the three selected public school adult education programs who rated each item in terms of the extent to which emphasis had been placed upon that area of adult need.

From an analysis of these data the investigator evaluated each of the three programs in terms of goals and adequacy of program in meeting needs and interests of adults enrolled. Approximately one-half of the report consisted of a descriptive presentation of the findings. Each of the three programs was treated separately in terms of three aspects:

1. The adequacy of the percentage of population served;
2. The adequacy of means used to realize the goals proposed by the study; and
3. The adequacy of the program emphasis placed by teachers and directors of the program upon the problem fields of adults as determined by a comparison of adult interest on similar items.

The following excerpts from the summary of findings are illustrative:

> With respect to the percentage of population served by the public school adult education programs being evaluated, this study shows:
>
> 1. The percentage of population being served by the Northeastern Oklahoma Agricultural and Mechanical adult program of Miami, Oklahoma, is thirty-two hundredths of one per cent. The program uses one adult educational center of the thirty-six available for adult learning activities.
> 2. The percentage of population served by the Stillwater, Oklahoma, public school adult education program is four per cent. The program uses five of the available eight educational centers for adult learning activities.
> 3. The percentage of the population served by the Tulsa, Oklahoma, public school adult education program is one per cent. The program uses seven of the fifty-six educational centers available for adult learning activities. . . .
>
> The summary of the means used in reaching the goals proposed by the study is noted in terms of inadequacy and is applicable to all three programs included in this study.

Goal I. To Create a Desire for an Equality of Educational Opportunity

1. Not all of the people concerned with adult problems are being used to identify the needs and interests of adults.
2. Not all people, agencies, and institutions concerned with adult problems are being used to determine needs and interests of adults.
3. The use of a greater percentage of the avenues for contacting and informing people of adult learning activities is not found in any one of the three programs studied.
4. The segments of population served are only a small sampling of the socio-economic groups found in the three communities studied. . . .

. . . .

Goal III. To Develop an Informed, Constructive Thinker

With respect to the channels used to pass on information to adults, this study reveals that: (1) The Miami, Oklahoma, program uses seven out of the thirty-two channels listed. (2) The Stillwater, Oklahoma, program employs fourteen out of the thirty-two means listed. (3) The Tulsa, Oklahoma, program utilizes fifteen out of the thirty-two listed. In terms of the adequacy of utilization of ways and means of getting the facts to people, this study indicates that this area can be investigated with profit by each of the adult programs included in this study. . . .

With respect to the comparison between the program emphasis and adult interest in the general areas of adult problems, this study shows that the Miami, Oklahoma, program has an adequate emphasis in two of the eight areas; the Stillwater, Oklahoma, program emphasis has an adequate emphasis in four out of the eight adult problem areas; and the Tulsa, Oklahoma, program emphasis has an adequate emphasis in three out of the eight adult problem areas. . . .

In comparison of program emphasis with adult interest in specific adult problems, this study reveals the following:

Miami, Oklahoma, Program

Out of the eighty problem program emphases

> assignments made by the Miami adult education staff, only twenty are in agreement with student interest demand. . . .
>
> Stillwater, Oklahoma, Program
>
> Out of eighty problem program emphases assignments made by the Stillwater adult education staff, only thirty-five are in agreement with student interest demand. . . .
>
> Tulsa, Oklahoma, Program
>
> Out of eighty problem program emphases assignments made by the Tulsa adult education staff, only twenty-four are in agreement with student interest demand.[1]

The conclusions drawn by the investigator may be stated as follows:

1. The adequacy of public school adult education programs can be determined.

2. It is possible to establish appropriate goals for adult education, and that it is possible to identify methods and procedures which are based upon accepted educational principles as a means by which these goals may be achieved.

3. It appears that adequacy of the program can be measured best in terms of the extent to which the program meets the needs of the adults enrolled.

The two foregoing studies by Cheskie and Timken are attempts to evaluate specified public school adult education programs. Such descriptive studies are of primary value only to the researcher and to those who have responsibility for the direction of a given program in that the reported findings offer specific leads for the improvement of the given program. However, in terms of quality of research there are outstanding weaknesses which are similar to those in many evaluation research studies. First, both studies are ex post facto attempts to evaluate a program for which no plan for evaluation was made at the inception of the given programs. Thus, evidently, from the reports, no precisely stated objectives had been formulated for either of the programs. Secondly, the specially constructed instruments which were used must be considered of poor quality since no reported effort had been made to establish validity or reliability of the instruments. Third, the primary evaluation data were obtained from participants who were enrolled during a specified time; hence, no evaluation was made over time. Thus, the findings of these studies are not generalizable to other public school adult education programs.

---

[1] Ibid., pp. 75-80.

Cheskie assembled and analyzed a vast amount of data. These data, however, represented only the opinions of those currently enrolled and of those representing the community. Thus, the sample was inadequate in terms of evaluating this program over time and for generalizing results of findings to future participants. Cheskie pointed out, however, that this study was limited to currently enrolled participants. No effort was made to establish the validity or reliability of the instruments used to obtain data, and no pretesting had been made.

The treatment of data was simply that of establishing relationships by inspection and findings were reported descriptively on the basis of frequencies and percentages. No statistical tests were employed; hence, there is no basis for generalizing results even to future participants who may have similar characteristics and who enroll in similar courses. The conclusions as drawn cannot be considered warranted as conclusions drawn from research, per se. Instead, they included descriptive statements as a summary of findings.

Timken, as pointed out in the analysis of his study, was inconsistent in terms of his stated purpose of the study. Thus, the research methodology was inappropriate for either purpose.

His stated first purpose, i.e., to establish a process by which public school adult education programs may be studied, is germaine only to a study which purposes to validate a process by which public school adult education programs may be evaluated. One may infer that his reason for taking three public school adult education programs was logical had he tested his proposed process and determined the extent to which the process was discriminating. Thus, if he had stayed with his first purpose he would have had a study.

He immediately departed from this purpose and proposed to determine the adequacy of three selected public school adult education programs. In terms of this purpose, the selection of three public school systems was logical. However, he made no claim, except by inference, that the three selected public school adult education programs were representative of those in Oklahoma or in the nation. He stated that each of the three public school systems were typical ones.

CHAPTER III

LIBERAL STUDIES BY INSTITUTIONS

The purpose of this chapter is to report the analysis and evaluation of the research reports which are related primarily to the liberal function of Adult Education. Seven studies were identified which have been conducted under the auspices of the specified institutions as follows:

1. University Extension 1
2. Public School Adult Education . . . 3
3. Great Books Foundation . . . . 3.

As in the foregoing chapter, studies conducted under the auspices of each of the three institutions are treated separately. First, a critical analysis is made of the studies after which a summary evaluation is presented of the study or group of studies.

University Extension

Hadlock conducted a study to determine if adults evinced a significant increase in ability to think critically after participating in an adult education discussion program.[1] The discussion program was "World Politics," prepared and published by the American Foundation for Political Education and sponsored by University Extension, University of California at Los Angeles. The primary objective of the Foundation was to develop the ability on the part of each participant to think critically and to make independent judgements in matters of current affairs. The hypothesis of the study was "that individuals did show a significant improvement in ability to think critically during participation in a discussion program of World Politics."[2]

The research design involved the comparison of members in an experimental group with those in a control group before and after the discussion program. Three criterion measures were used: (1) Test of Critical Thinking, developed by the Cooperative Study of Evaluation in General Education, (2) The Cooperative English Test, Reading: Level of Comprehension and Vocabulary, and (3) Sanford's and Older's Short Authoritarian Scale derived from the Berkley F(ascism)

---

[1] Alton Parker Hadlock, "A Study of the Development of Critical Thinking through Adult Discussion Groups" (unpublished Ed.D. dissertation, University of California, Los Angeles, 1958).

[2] *Ibid.*, p. 49.

Scale. Because only one form of the critical thinking test and the authoritarian scale was available, identical forms were used in both administrations; however, a different form of the reading test was administered. In addition to the purpose of comparing changes, the critical thinking test was administered to members of the control group in order to ascertain the practice effect involved in taking the same test twice. The reasons given for including the reading tests were: (1) to determine if a significant relationship existed between the reading ability of the participants and their scores on the critical thinking tests, (2) to provide a control by which it could be determined if an increase in critical thinking ability was associated with a corresponding increase in reading ability, and (3) to afford a means by which participants could be compared with college students who had taken the same test previously. The primary purpose for including the authoritarian scale was to determine if a significant difference existed between those participants who were authoritarian and those who were not authoritarian.

Participants in the study included fifty-seven adults who volunteered from six different discussion groups in World Politics. This number represented 63 per cent of those who completed the course in the six groups studied. Of the original fifty-seven who completed the pretests, thirty-eight completed the posttests. Thus, the experimental group consisted of thirty-eight members. The control group consisted of adults who either were members of the faculty, spouses of faculty members, or acquaintances of the faculty. They were volunteers also. Selection was made on the bases of educational level and of current non-participation in an organized educational endeavor. All members of the control group had earned at least a baccalaureate degree, and no person was a current participant in an organized educational program.

The test data were analyzed statistically to answer the following proposed questions:

1. Was the experimental group an adequate sampling?
2. How did the reading scores, both the level of comprehension and vocabulary, of the experimental group rank with the reading scores of college senior in liberal art colleges?
3. Did practice in taking the same pretest and posttest of critical thinking affect the final test score?
4. Was there a significant difference in the change test scores of the experimental group from the change test scores of a control group in critical thinking?
5. Was there a significant change between pretest and posttest scores in critical thinking for the experimental group?
6. Was there a significant change between the pretest and posttest scores in critical thinking for the

control group?

7. Was there a significant change between the pretest and posttest scores in the level of comprehension for the experimental group?
8. Was there a significant change between the pretest and posttest scores in vocabulary for the experimental group?
9. Was there a significant difference in the mean change scores in critical thinking between the experimental and the control groups?
10. Was there a significant difference between the pretest and posttest scores of the most authoritarian and the least authoritarian individuals on the critical thinking test?
11. Was there a significant difference between the means of the six different discussion groups on the critical thinking test?
12. Was there a significant difference between the change test scores of groups of individuals with varying amounts of education for the critical thinking test?
13. Was there a significant difference in the mean change scores of critical thinking between the five different age groups?[1]

Essentially, in the analyses of these data, each question was cast as a null hypothesis and the difference between the mean test scores, or the difference between the differences of mean scores, or both, were tested for significance. T-tests or F-tests were applied, and the .05 or higher level of confidence was accepted for significance. However, in one case a difference between pre- and posttest scores was recorded at the .10 level of confidence. The analyses which were made to answer questions one, two, three, and twelve are reported briefly in the subsequent paragraphs in order to illustrate the statistics employed.

---

[1] *Ibid.*, pp. 9-10.

To determine if the experimental group was an adequate sampling of the original fifty-seven, three analyses were made: (1) the mean of the critical thinking pretest scores for the experimental group was compared with the mean score for the original fifty-seven, (2) the mean of the pretest scores on critical thinking for the experimental group was compared with the mean for the nineteen who took the pretest but who did not take the posttest, and (3) the mean of the pretest scores on critical thinking for the nineteen was compared with the corresponding score for the original fifty-seven. Since no significant difference was found to exist in either of the three statistical analyses, it was concluded that the experimental group of thirty-eight participants was an adequate sample of the original fifty-seven.

To compare the experimental group with 50,000 college seniors in typical liberal arts colleges, the reading comprehension scores for participants in the experimental group were listed opposite the percentile ranks of college seniors who earned the same scores. The percentile ranks of the experimental group were listed also for comparison.

The mean and standard deviation for the college seniors were 65.2 and 10.3, respectively, and the mean and standard deviation for the experimental group was 65.2 and 9.18, respectively. Thus, the similarity of these two distributions was evidence that the experimental group was a select group in that it did not differ significantly from college seniors. Similarly, comparison was made for the vocabulary pretest scores. The mean and standard deviation for the seniors were 63.8 and 9.7, respectively, and for the experimental group the mean and standard deviation were 72.6 and 6.9, respectively. Accordingly, the investigator concluded that it was evident on the average that the participants in the discussion programs possessed a larger vocabulary than did the college seniors since the percentile ranks showed that their scores were higher throughout the entire range.

To determine if taking the same pretest and posttest of critical thinking affected the final test score, three analyses were made. First, the means of pretest and posttest scores for each group, experimental and control, were tested for significance of difference. For the experimental group the difference was found to be significant at .001 level of confidence. For the control group, the difference was recorded at .10 level of confidence. Then, comparison was made of the "change scores" for each group; i.e., the significance of the differences between the difference of the means was tested. The investigator pointed out that if no difference was found, the two groups would be considered to be similar, and the differences in scores would be attributed to extraneous variables or to practice in taking the same test twice. But if a significant difference were found to exist, then the difference would be attributed to change. For this comparison, the t-score was found to be 3.40 which was significant at the .01 level of confidence. Thus, it was concluded that "a significant difference did exist between the two groups, and that the increase in scores of the experimental group was not due to practice effect or to extraneous variables."[1]

To determine if educational background had a significant affect in producing changes in ability to think critically, the following analyses were made. Participants were divided into four groups according to their level of education. Group I embodied those individuals who had less than two years of college; Group II consisted of those with two years of college and more, but those who had not earned a baccalaureate degree; Group III included those who had earned a college degree and some who had engaged in graduate work, but who had not completed a graduate or professional degree, and Group IV involved only those with graduate or professional degrees beyond the baccalaureate. An analysis of variance was employed to measure the differences between the means and the mean of the entire group for the change critical thinking test scores. The F-test revealed a ratio of 1.62 which was too small to warrant the rejection of the null hypothesis. Thus, it was concluded that the means of the groups with different levels of education did not vary significantly from the mean of the entire group. However, it was observed that the mean of Group IV was greater than that for the other three groups; and, the mean of the change scores for each group had increased with the level of education. Therefore, Group I, the group possessing the least amount of educational background, was compared with Group IV, the group possessing the greatest amount of educational background. Again, the analysis of variance was employed, and the F-ratio was found to be 4.81 which was large enough to warrant rejection of the null hypothesis at the .05 level of confidence. Hence, it was concluded that adults who had had the greatest amount of education made a significantly greater increase in critical thinking than did those who had had the least amount of education.

---

[1] Ibid., p. 63.

In like manner the statistical analyses were reported clearly for each question. Several tables presenting the data were included in the appendices. It was pointed out that conclusions reached in this study could be generalized only to the participants in the discussion groups, and not necessarily to the population as a whole. The twelve conclusions drawn as a result of the analyses of the data may be summarized as follows:

1. The experimental group of thirty-eight was an adequate sample of the group of fifty-eight who volunteered for the research project.

2. The experimental group was a select group in terms of both reading level of comprehension and vocabulary, as determined from the comparison made with a group of seniors in liberal arts colleges.

3. A significant difference existed between the increased critical thinking test scores of the experimental and control groups at the .05 level of confidence.

4. A significant increase in the change critical thinking test scores existed at the .001 level of confidence for the experimental group.

5. A significant increase was recorded in the change critical thinking test scores at the .10 level of confidence for the control group.

6. No significant increase was made in the reading level of comprehension; however, a significant increase occurred in the change scores of the

vocabulary tests at the .001 level.

7. No significant difference existed between the change critical thinking test scores for those who were classified as most authoritarian and those classified as least authoritarian.

8. No significant difference was found between the means of the change critical thinking test scores for the six different groups.

9. Difference between the change scores of critical thinking for groups with varying levels of educational background was not significant; however, a significant difference was found between the means of the increased critical thinking scores of the group with the least amount of education and the group with the greatest amount of education.

The method employed by Hadlock was appropriate. His study was neatly designed and reported in a consistent, clear manner. The use of a control group enabled him to state that reported changes in the participants were a direct result of the treatment. The sample may be considered an adequate sample of the portion of the classes who initially volunteered. Since he used only members of the discussion groups who volunteered, one cannot be certain that the same findings would have resulted from a random sample of all participants. However, since participation in such discussion groups is wholly voluntary, the researcher probably had no other choice in the selection of a sample. The thirty-eight participants in the study did represent approximately 42 per cent of those who completed the course in the six groups studied, and, as a group, they met the criteria necessary for considering them a random sample of the fifty-seven who originally volunteered.

The measures used were of high quality and appropriate for stated purposes. The fact that Hadlock stated precisely his purposes for the administration of each test and analyzed data accordingly enhanced the quality of his research. Statistical procedures employed were appropriate and results obtained from each analysis were reported and interpreted clearly. The .05 and higher levels of confidence were used for designating significance. At one point the .10 level of confidence was recorded, but further analysis indicated that the investigator did not consider this to represent a true change.

The conclusions drawn were sound and warranted in terms of reported findings. Furthermore, Hadlock pointed out that the conclusions reached in this study could be generalized only to the participants in these discussion groups and not necessarily to the population as a whole. The study contributes to practice. It offers an acceptable method for evaluating other adult education discussion groups which have similar purposes. However, a random sample of subjects should be used if results are to represent those of the population.

Public School Adult Education

Of the studies conducted under the auspices of Public School Adult Education, three relate to the liberal function of Adult Education. Since the three studies differ both in terms of content and of research methodology, each is reported. An analysis of the studies is made, after which a summary evaluation of the three studies is presented.

Burkey examined the role that the creative arts were playing in the lives of adult students.[1] Purporting that participation in the arts was an important facet of adult education and that it met a variety of needs for a large number of adults, Burkey attempted to identify and clarify those needs. Essentially, the study was a quest for a body of knowledge to assist the professional personnel of adult education in the development of a program of creative arts designed to meet the specific and personal requirements of the mature individual.

---

[1] Robert Lyle Burkey, "Student Responses to an Evening Session Public School Adult Education Art Program" (unpublished Ed.D. Type C Project, Teachers College, Columbia University, 1961).

Before approaching the problem of constructing an instrument to secure data, Burkey reviewed the literature; visited a number of adult art classes in session, and talked with administrators, teachers, and participants in order to develop a "feel and knowledge of teaching methods, course content, and student interests." After tentative areas of concern were identified, he held informal interviews with actively participating adults to test the suitability of the areas. Thusly, the framework of the questionnaire evolved into three major sections: (1) why adults come, (2) what classroom procedures they prefer, and (3) what happens to the individual as a result of having attended the art classes. Two minor sections of the instrument elicited information as to age, sex, formal education, occupation, and previous adult art class experience.

Burkey then proceeded to develop additional divisions for each of the major sections. He formulated a rationale both for the divisions and for the items constructed within each of the divisions. Briefly, the subsequent paragraphs report this process.

Five general motives for attendance became apparent: (1) the social motive, the need to be a part of a group having common interests, (2) the improvement of vocational skills, (3) recreational or therapeutic values that might be derived from the experience, (4) pre-knowledge that the class and/or the teacher had a great deal to offer, and (5) the appeal of the art activity itself. Four items, ranging from the general to the specific, were designed for the social motive. These were: (1) "I had enjoyed a previous group so much that I wanted to join another," (2) "I wanted to meet new and different people," (3) "I wanted to meet people with interests similar to my own," and (4) "I wanted to be with people." Two statements were developed to cover the possibility that participation in the art program might be to improve vocational skills. One statement was directed toward the middle and older age groups, and one toward the younger adults who might be thinking ahead to a specific vocation.

In an attempt to discover how great an influence was exerted upon initial motivation by recreational or therapeutic values, three items at three different levels were used. First, to question the need to fill leisure time, "I wanted to do something with my leisure time," was used. Secondly, designed to include those people who simply came to relax, the item was, "I wanted a place to come for relaxation." The third item, "I wanted to get away from my everyday life," expressed the essence of many comments made during the early interviews, and embodied some elements of a desire to escape the problems of everyday living through absorption in the companionships or activities offered by the adult art classes.

Two items, "The teacher was recommended to me," and "A friend of mine told me about the class," were used to examine the importance of comments by friends about a teacher's appeal and/or the excellence of a particular course as they specifically related to classes in the creative arts. Five remaining items were designed to explore the relative importance of the art activity itself. They related to knowledge about art, to the facilities, to tools and materials, and to techniques. The five items were: (1) "This type of art work was new to me and I wished to know something about it," (2) "I was familiar with this type of art work but wished to know more about it," (3) "I wanted a place to do art work that I could not do at home," (4) "I wanted to develop competence with tools and materials," and (5) "I wanted to explore new art materials and techniques." Item No. 2 above was used because many participants had attended the same course for a period of years but still expressed the desire to learn more about it, while other participants had sampled every art class offered in the Center to extend their scope of knowledge about the arts. Respondents checked one of four choices of degree for each of the items in the first major section: "very much," "some," "little," or "does not apply."

The second section of the instrument was designed to determine the degree to which participants preferred each of seven classroom procedures. Beginning with a procedure that was highly controlled and dominated by the teacher, each item sequentially became less teacher-centered and more student-centered. The items were: (1) "Teacher's lecture to the entire class," (2) "Teacher's demonstration of technique," (3) "Individual instruction from the teacher," (4) "Discussion with fellow students," (5) "Viewing other students' work," (6) "Reading about art and art techniques," and (7) "Working alone in the classroom." Respondents checked one of four choices for each of these items: "very much," "some," "little," or "no help."

The last major section was designed to get some index of the increase in specified outside activities occurring as a result of participation in the public school adult art program. Fifteen items, concerning five different types of activities, were included. Since a majority of participants interviewed had indicated that they painted, sculptured, or worked in other media at home whenever they could arrange time, an important item in this section was "Being stimulated to do art work at home." Five items were used to ascertain if personal involvement in the arts had increased as a result of art classes. These items were: (1) "Attending museums of art," (2) "Attending art exhibits," (3)"Reading about art and art techniques," (4) "Buying reproductions of works of art," and (5) "Buying original works of art." Three items were developed to determine transfer of interest from one art form to another: "Attending the concert," "Attending the theater (stage plays)," and "Attending the ballet." The six remaining items were used to determine how extensively participation in the art courses had affected changes in personality and self-concept. These were: (1) "Helping other people (outside the class) to make decisions involving artistic judgement," (2) "Gaining confidence in my own abilities," (3) "Seeing myself in a new light," (4) "Feeling more alive," (5) "Becoming more critically aware of the forms and relationships of form and color that surround me in my daily life." For each of the items in section three the respondent checked either "increased," "no change," or "decreased."

In addition, the instrument elicited suggestions for improving the adult art classes, and one-half of a page was left blank for this purpose. The first major section and the two minor sections were to be answered by all respondents, but only those with previous experience in adult art classes were to respond to the second and third major sections, classroom procedures and changes in specified outside activities, respectively.

Burkey selected for his sample two adult education centers which served a cross-section of adults enrolled in the public school adult education program of New York City. The Taft Youth and Adult Center in the Bronx was chosen because of its comprehensiveness of course offerings and of its location in a residential, middle socioeconomic setting. Of the 219 adult education courses offered, 22 were art courses. The Wagner Youth and Adult Center in Manhattan, the newest of the public school centers, was located in a low socioeconomic area and offered a more restricted program. Of its fifty-eight courses, eight art courses were offered.

For a pre-test, the questionnaire was administered in another adult education center to fifty people enrolled in an art program similar to that offered by Taft and Wagner.

Minor revisions were made after which Burkey personally administered the questionnaire, during the regular class periods, to the 438 adults enrolled in the thirty classes. The administrations were consistent in that Burkey moved from class to class; gave identical instructions; passed out the forms; remained in the room as the respondents completed the questionnaire; answered any questions that arose, and collected the forms upon their completion. In each case, this administration required from twelve to fourteen minutes. After collecting the forms, he conducted a brief discussion pertaining to such matters as the disposition of the results, strengths and weaknesses of the existing art program, and to general questions in relation to the study. He noted in the report that many verbal suggestions and comments made at this time yielded a wider range of suggestions from participants than was indicated in written form. Of the 400 respondents who completed forms that were usable, 105 were males and 295 were females.

The data were tabulated in ten groupings, males in the five age groups (18-29, 30-39, 40-49, 50-59, 60 and older), and females in the five age groups, and reported in frequencies and in percentages. Also, the sample was divided into two parts: those respondents who were new to the program, and those who had had previous adult art education experience. Of the 400 respondents, 68.7 per cent, or 271, were experienced participants. The major section of the report consisted of a descriptive interpretation of the data obtained on each item of the questionnaire. Fourteen tables presenting the data in frequencies or in percentages were interspersed within the content. Major findings may be summarized briefly as follows:

> The most influential motive operating in art class attendance was the desire to participate in the art activity. The social motive had some importance for the older adults; the desire to improve or develop vocational skills was almost a negligible factor; and recreation and therapeutic values were important to a number of participants, especially older ones.
>
> Preferences of classroom procedures indicated clearly that close supervision by the teacher was very important to these adults. Teacher demonstration of techniques was the most preferred classroom procedure with individual instruction and lecture following closely.
>
> One of the greatest changes in the behavior of the participants outside of class was in pursuing their art work at home. A remarkable result of the adult art classes was the increase in outside activities, especially those which had the closest relationship to art classroom activities.
>
> The aspects of self-confidence, self-image, and intensification of living that might be gained through the art experience received a higher percentage of response to "increased" than did any of the other items in the third major section.

Conclusions as such were not drawn; rather, two lengthy chapters were devoted to "Implications and Recommendations" and to "Summary," respectively. Of the many implications pertaining to the teaching approach, to course content, and to curriculum considerations, the following are illustrative.

> The academican, or the leader of a 'school' of art has no place as an adult art educator. . . .[1]
>
> The art teacher of adults must do the greatest part of his teaching on the individual level. . . .[2]
>
> He must be prepared to serve as art teacher, art critic, and art historian to meet the adult's hunger for knowledge about the arts . . . .[3]
>
> The suggestions that have been made should be considered as supplementary to the present program and not meant to supplant them. . . .[4]
>
> To insure continued satisfaction, courses of instruction should be provided for students at several levels of artistic maturity. . . .[5]
>
> The art programs at Taft and Wagner are very good. They offer a wide variety of classes and serve a great many adults. Basic weaknesses are not caused by specific programs but are inherent in the present role of art in adult education.[6]

Matherne analyzed and evaluated an experimental program in basic economics which was designed to increase the level of economic understanding among adults.[1] He stated: "The hypothesis to be proved is that the level of economic understanding can be raised significantly through the use of selected methods, materials, and other resources available in most communities."[2] An analysis was made of the following instructional methods: (1) lay persons specially trained for the program, (2) resource people, (3) basic economic materials, (4) visual aids, and (5) conference leadership techniques. The criterion measure used to evaluate the program was a test which was administered to subjects before and after participation in the program. He pointed out that the evaluation was to be made only in terms of the five teaching methods. He also proposed that the hypothesis would be established "by using the teaching method which will be described, and by measuring the level of economic understanding before application of the teaching method and after the application of the method."[3]

---

[1] *Ibid*., p. 69.

[2] *Ibid*.

[3] *Ibid*., p. 72.

[4] *Ibid*., p. 77.

[5] *Ibid*., p. 80.

[6] *Ibid*., p. 83.

The program, sponsored by the Mott Adult Education

---

[1] Allen James Matherne, "An Analysis and Evaluation of an Experimental Program Designed to Increase the Level of Economic Understanding among Adults" (unpublished Ph.D. dissertation, Michigan State University, 1960).

[2] _Ibid_., p. 24.

[3] _Ibid_., p. 56.

Program, in cooperation with the Industrial Relations Center and the A C Spark Plug Division of General Motors Corporation, was conducted in Flint, Michigan, during the fall of 1959. The population for the study was presumed to be the adults who resided in the city and in its environs. An attempt was made to include a representative group of participants in each of the thirty-seven community schools in the City.

Adults were selected and invited to participate either by the building principal or by the director of community activities. They were selected on the basis of: (1) past participation in the general program of the community school, (2) prior enrollment in adult education classes offered at the school, (3) participation in one or more school organizations, and (4) presumed interest in the area of economics. Matherne pointed out that the persons who were involved probably differed in characteristics from those who would have been attracted by a general announcement. A total of 841 adults in thirty-seven community schools enrolled in the basic economics course which consisted of eleven sessions of two hours duration. Registrations for each class ranged from eleven to thirty-three.

The teaching staff, or discussion leaders, were recruited from lay persons normally found in most communities. Young men in various middle management positions of the A C Spark Plug Division, General Motors Corporation, were selected. They were trained in conference leadership techniques and in basic economics. The conference leadership training materials developed by the Center were used in the training sessions for the leaders. Personnel from the Center assisted in the training sessions which included a series of eight seven-hour sessions. These sessions were devoted to lectures, demonstrations, and practice in the leadership function. The materials used in the adult classes had been developed by the Industrial Relations Center of the University of Chicago. They were designed for use with high school students and with adult groups, and had been pretested in the high schools of Flint. The text material consisted of eleven topics, each of which was composed of several sub-topics. The general topics were: (1) "A Look at Our Economy," (2) "Managing Your Income," (3) "Competitive Prices in Action," (4) "Producing for Better Living," (5) "Capital--Key to Progress," (6) "Your Wages," (7) "Profits--Sparkplug of the Economy," (8) "Understanding Money and Banking," (9) "Ups and Downs of Business," (10) "Security and Opportunity," and (11) "America and the

World Economy." A list of supplementary reading materials was prepared for the use of participants. These materials were required reading for the leaders. Various kinds of visual aids, including both those which were carefully selected and those which were specially prepared by the Center, were used during class sessions. Many ideas were illustrated graphically on "flok cards" which either were displayed on flannel boards or on "cells" and shown on an overhead projector. Each class session called for the employment of from eight to twenty-nine "flok cards" and from three to eleven "cells."

A number of resource people, including bankers, economists, and farmers, were invited to class sessions at appropriate times. They had been selected previously and had agreed to participate when requested. Several group techniques were employed to disseminate the specialized types of information which these resource people possessed. The resource person either lectured; replied to questions asked by individual participants; served as a member of the group and advanced technical information at appropriate times; or responded to questions submitted by a panel composed of group members.

The test, developed previously by the Center, consisted of fifty true-false statements designed to test understanding in the eleven areas represented in the text materials. This test was administered to subjects at the first meeting of the class, and readministered at the last class session. However, it was omitted at the discretion of the group leader, if he felt the test posed a threat or might be a source of embarrassment to subjects. Consequently, complete test results were returned for thirty of the thirty-seven groups. Of the 841 enrolled adults, 650 completed the pretest and 462 completed the posttest.

Test data were analyzed in terms of subjects' performance on both tests by statements and by topics. Matherne pointed out that since fewer people took the posttest than did the pretest, that results were expressed both in numbers and percentages, with increases or decreases expressed in percentage points. To illustrate: For statement one, 232 of the 650 subjects gave correct responses on the pretest, and 255 of the 462 responded correctly on the posttest. Thus, 35.8 per cent of the subjects gave correct answers on the pretest and 55.2 per cent responded correctly on the item on the posttest. Consequently, for statement one, the increase, or growth, was 19.4 percentage points.

Program-wide results from this analysis were summarized as follows:

> The range of correct responses on the pretest was from (101/650), or 15.6%, to (631/650) or 97.1%. The range of scores on the posttest was from (107/462), or 23.2%, to (450/462), or 97.4%.
>
> The increase in percentage points from pre to post ranged from a low of minus 2.5% to 24.9%.[1]

Similarly, an analysis by topics revealed the percentages of correct responses on both tests. These findings were summarized as follows:

> The range of numbers and percentages of correct responses on the pretest was from a low of (349.4/650) or 53.8%, to a high of (576.5/650), or 88.7%. On [the posttest the numbers and percentages of correct responses ranged from] (280.0/462), or 60.6%, [to] the high (399.0/464), or 91.5%.
>
> The increase of post over pre, expressed in percentage points, was from a low of 2.5% to a high of 12.4%.[2]

After making six recommendations pertaining to the improvement of programs in economic education for adults, the investigator stated:

> It is acknowledged that the experiment analyzed and evaluated during the course of this study represents but a small segment of the total problem of raising the level of economic understanding among adults. Attention might be directed to other programs in economic education for adults, involving selection of different populations, employment of different methods, and utilization of different materials. Thus, a body of knowledge about this area should be developed which could prove useful to the adult educator.[3]

---

[1] *Ibid.*, p. 105. [2] *Ibid.*, p. 106. [3] *Ibid.*, p. 110.

Shrock investigated certain behavioral outcomes of a selected adult education course in ceramics.[1] The stated problem was "to determine whether the methodology employed in teaching adult education ceramics classes is helpful in bringing about increased personal adjustment in the participants and the degree to which this is accomplished as measured by certain standardized and non-standardized self-report instruments."[2] The research methodology employed was that of testing the following null hypotheses:

> General Hypothesis.--There is no change in the personal adjustment of individuals participating in adult education ceramics courses as a result of teaching methodology or as a result of the experience with the craft methods and materials.
> Hypothesis 1.--No significant difference in adjustment exists between two groups of students at the termination of the experience in which each group is exposed to methodology differing from the other. The areas of adjustment considered are those measured by the Guilford Inventory of Factors STDCR.
> Hypothesis 2.--No significant difference in adjustment exists within each group between the initiation and the termination of the experience as measured by the Guilford Inventory of Factors STDCR.
> Hypothesis 3.--Any positive or negative changes of significance in adjustment are related to the differential methodology used by the teachers.[3]

In addition, answers to the following questions were sought:

---

[1] John G. Shrock, "Functions of Methodology in Adult Education Ceramics Courses" (unpublished Ed.D. dissertation, Stanford University, 1953).

[2] Ibid., p. 26.

[3] Ibid., pp. 26-27.

1. What are the main referral sources for an adult education ceramics course?
2. What are the stated reasons for taking a course of this nature?
3. How much leisure time do participants in this activity report?
4. What kinds of hobbies do the participants have?
5. What is the extent of adult education experience of the students?
6. What are the students expected outcomes of the course?
7. What are the reported final outcomes of the course?
8. Did the experience tend to raise or lower the subjects' concept of their artistic ability and appreciation?
9. Do the participants as a group feel that ceramics is a relaxing activity?[1]

Upon listing the above questions the investigator pointed out that the data from which the tentative answers were derived were not handled quantitatively in a statistical sense.

The study was conducted in an adult education center in which approximately 100 adults were enrolled in four ceramics classes taught by two teachers. The duration of these classes was approximately twelve weeks, and classes met for a period of three hours twice a week. Subjects in the study were class members who volunteered to participate. Sixty-six women completed all pre- and post-tests. The two groups were distinguished on the basis of the teacher under whom they studied: Teacher A and Teacher B. Group A, the experimental group consisting of thirty-four participants, was taught by Teacher A; Group B, the control group consisting of thirty-two members, was taught by Teacher B. Groups were considered to be homogeneous since no significant differences were found between the two groups on five variables: age, marital status, number of children, and educational level, and sex.

---

[1] _Ibid_., pp. 27-28.

Teacher A was identified as the "artist--craftsman." He emphasized free expression, a minimum of discipline, prohibited copying, refused to do the students' work, and maintained a goal of unique products for each adult. Teacher B, identified as a "teacher--craftsman," did not consider herself a creative artist but one who was interested in people. She emphasized therapy, high standards, discipline in terms of method and form, the study of models, and condoned imitation of good pieces. The nature, content and technical aspects of the two classes were fundamentally the same. Hence, the prime variable in this experiment was the difference between the goals of the teachers and the ways in which the teachers related to their students in order to reach their goals.

Three instruments were used as criterion measures.

The Kuder Preference Record-Personal was administered before the class experience, but it was not used at the termination of the classes. This test was eliminated because many of the participants, especially those over sixty years of age, complained of the difficulty encountered in reading the small print and in responding to items because of the smallness of the pins and spaces. An Inventory of Factors STDCR by J. P. Guilford was administered both before and after the experience. The same form was used for both administrations. Questionnaires were designed to obtain descriptive information and to secure responses of an attitudinal and self-evaluative nature. The pre-experience questionnaire consisted of fourteen items and the post-experience embodied ten items. Copies were included in the appendix of the report. All tests were self-administered at home because the policy of the public school system prohibited the administration of tests during class periods.

All test data were analyzed statistically. Three types of comparisons were made: (1) inter-group comparisons of results obtained from the administration of the Kuder Preference Record-Personal Test, (2) intra-group comparisons of results obtained from the two administrations of the Guilford Inventory of Factors STDCR, and (3) inter-group comparisons of results obtained from the two administrations of the Guilford Inventory of Factors STDCR. Each is reported briefly in the subsequent paragraphs.

1. Intergroup Comparisons: Kuder Preference Record-Personal.--Results from the one administration of this test were analyzed to compare the preferences of the two groups. Four variables, age, marital status, number of children, and educational level, were used for matching pairs of subjects. All participants in the study were females. In order to arrive at an equal number of subjects in each sample, two members of Group A were eliminated on the basis of extremes in age, sixty-two and sixty-eight years. Shrock, using a band of five years, was able to match eight pairs of subjects on four of the variables, eighteen pairs on three, and six pairs on two of the variables. Thus, the computations were based on thirty-two pairs of subjects matched on two or more variables. The significance of the difference of the mean differences between paired scores was computed for each of the five scales. He found no statistically significant difference for any of the five scales. However, for scales A, C, and D, Group B was slightly higher than was Group A, and for scales B and E, Group A had higher mean scores than did Group B. Since the largest differential

occurred for scale D favoring Group B, Shrock pointed out that Group B showed slightly more preference for pleasant and smooth personal relations which are free from conflict than did Group A. He stated: "In view of these findings it is necessary to accept the null hypothesis that there is no difference between the paremeters in question for these two groups."[1]

2. Intra-group Comparisons: Guilford Inventory of Factors STDCR.--Different pairs of subjects were matched for each of the five scales. The variables used sequentially for matching were: (1) sex, (2) scores on the pre-test, (3) age, and (4) educational level. The same statistical procedures were employed to determine if a significant change had occurred within the subjects over the period of time on the factors measured by this instrument. Shrock found that for Group A there was a gain in a positive direction for each of the five factors. However, for only two factors, D and C, were the differences in gains statistically significant at the .05 and .01 levels of confidence, respectively. He pointed out that Group A was significantly less depressed and more emotionally stable at the close of

---

[1] *Ibid*., p. 65.

the treatment. Even though there were changes in a positive direction for Group B, no statistically significant difference in change existed on any factor.

3. Inter-group Comparisons: Guilford Inventory of Factors STDCR.--The differences between the differences of pretest and posttest scores for the two groups were examined. Findings indicated for all factors that Group A was somewhat better adjusted than was Group B. However, Group A significantly exceeded Group B only on one factor, emotional stability (.01).

The findings from the above three comparisons may be summarized as follows:

1. Significant gains in feelings of cheerfulness and optimism were found for Group A while Group B remained the same.
2. Subjects in Group A made significant gains in emotional stability. No change was found in subjects of Group B in this respect.
3. After treatment, Group A was significantly more emotionally stable than was Group B.

Thus,

Null hypothesis 1 was accepted for Factors S, T, D, and R. It was rejected for Factor C.

Null hypothesis 2 was accepted for Group B. For Group A, it was accepted for Factors S, T, and R and rejected for Factors D (depression) and C (emotional stability).

Null hypothesis 3 was accepted.

The results from the analysis of the responses to the questionnaire are summarized below. Each questionnaire stem is briefed and the first and second most frequently given responses are reported.

1. Referral sources
   a. Self-referred
   b. Referred by friend
2. Reasons for taking course
   a. To make useful objects for themselves and for gifts
   b. To meet people and make friends
3. Leisure time reported by subjects: An average of 2.75 hours per day
4. Kinds of hobbies
   a. Gardening
   b. Sewing
5. Extent of adult education experience: An average number of 1.74 courses taken previously
6. Subjects' expected outcomes
   a. Increased skills in the crafts
   b. Recreation and relaxation
7. Subjects' reported outcomes
   a. Technical skills
   b. Increased self-understanding
8. Majority of subjects indicated improved feelings, and
9. Majority of subjects indicated that activity was most relaxing.

After presenting these findings, the investigator stated: "It may be possible to conclude that the above findings show that in this course students do realize more satisfactory personal adjustment as a result of either the teaching methodology or as a function of the activity."[1] He also pointed out that all conclusions reached in this study are to be considered tentative and that they are not amenable to broad generalizations because of: (1) the exploratory nature of the study, (2) the small number of subjects involved in the experiment, and (3) the limited number of measuring instruments used to obtain data.

Of the three foregoing evaluation studies conducted under the auspices of Public School Adult Education,[2] only the methodology employed by Shrock may be considered appropriate. The two studies reported by Burkey and by Matherne, respectively, were descriptive in nature. In fact, they were projects rather than evaluation studies, per se. The objectives of Burkey's project either could not or were not precisely defined; thus, they could not be measured to any refined degree. The objectives of the project reported by Matherne involved the use of certain methods and resources as well as the cognitive objectives which were implied in the test. Thus, he made an analysis of the project in terms of the specified methods and resources, and evaluated it in terms of percentage points gained on each statement of the test and on topics studied. Data for these two studies were reported in frequencies and percentages, and no efforts were made to generalize results.

---

[1] Ibid., p. 86.

[2] Supra: Burkey, p. 223; Matherne, p. 232; Shrock, p. 240.

In all three investigations the subjects who were studied were chosen on the bases of their availability and of their willingness to cooperate in the respective studies. Burkey selected two adult education centers which he reported were representative of the public school system. He obtained data on 400 of the 438 enrollees in art classes at the two centers. Subjects studied by Matherne and Shrock were volunteers. Matherne exerted no special effort to obtain test results on all participants. Because of the way in which data were analyzed, one may suspect that there were subjects who responded only to the first or to the second administration of the test. Shrock's sample was limited to those who volunteered, 66 subjects out of a possible 100 or more. Hence, one suspects that results were biased since it is not known how the 34 nonvolunteers may have reacted to the treatment. It must be recognized that samples selected on volunteer bases cannot be considered as representative of the population.

Shrock was the only researcher who used a control group. Both Burkey and Matherne used a single sample. Hence, they are not justified in implying that their obtained results were a direct result of the treatment since no control groups were utilized.

The quality of measures used in the three studies varies considerably. Burkey constructed an instrument but made no effort to establish its validity or reliability. His method of analyzing data was appropriate for scales in which the units may not be presumed to represent equal distances. Matherne twice administered a fifty-item true-false test which had been designed by the Industrial Relations Center of the University of Chicago. Validity of the test may be assumed since it was developed to test the specified content. However, Matherne did not report any information on its reliability or validity. He did report that the materials used had been tested in local high schools. Hence, it may be assumed that the authors established reliability for the test, but one may question if such reliability holds for adults. Two of the three instruments used by Shrock may be considered of high quality. However, the Kuder Preference Record-Personal test evidently was not an appropriate type for his purposes since the second administration was eliminated because of complaints, especially from the older adults. It does not appear that the use of the questionnaire to secure responses of an attitudinal and self-evaluation nature added anything to Shrock's study.

The statistical procedures employed by Shrock were both sound and appropriate for his purposes. The .05 and higher levels of confidence were accepted for designating significance. Neither Burkey nor Matherne treated his data statistically; instead, each reported them in frequencies and in percentages. Thus, no confidence levels were used.

The conclusions as drawn from the three studies vary in terms of soundness. Burkey drew no conclusions as such; rather, he reported descriptively the results derived from the immediate responses given by participants. Matherne also reported findings in a descriptive manner. He implied that his stated hypothesis was proven by the resulting effects of the experimental program in basic economics. In terms of his research design, even the stating of an hypothesis is open to serious questioning. Thus, it must be recognized that the studies by Burkey and by Matherne were not appropriately designed to justify conclusions. The conclusions as drawn by Shrock were sound in terms of reported findings. He did what he proposed to do and reported findings accordingly. Furthermore, he placed limitations on his conclusions and gave three reasons why they must be considered tentative and not amenable to broad generalizations. However, the principal limitation was the sample used and the way in which it was selected which he did not specify as a limitation.

All three studies contribute to practice to some degree. The study by Burkey contributes to practice only in that it offers suggestions for the improvement of adult education art programs. His ideas, however tentative, may be tested in other situations. Matherne's study probably contributed to immediate practice in the setting in which it was conducted. It offers a good starting point for other adult educators interested in the evaluation of programs in economic education. The study by Shrock contributes to immediate practice in that it may be duplicated by other researchers for specific programs. It may be possible that his work will contribute to a general theory if it is utilized with other similar studies as a basis for theory development.

The Great Books Foundation

The Great Books Foundation, a non-profit organization, offers a national program for the liberal education of adults. It stimulates the formation of study-discussion groups and provides readings and guides for self-study and discussion. The readings are organized sequentially into blocks of one year in duration; thus, theoretically, books and materials should be read and studied in sequence. The Great Books Program, originally developed by Mortimer Adler and Robert E. Hutchins, is well known both nationally and internationally. For example, in 1959-60, the Great Books Foundation enrolled 50,000 participants in 2,700 groups in more than a thousand communities in the United States and abroad.

The Fund for Adult Education, established in 1951 by the Ford Foundation, laid particular emphasis upon study-discussion programs in the liberal arts. This was done not only by conducting its own Experimental Discussion Project, 1951-1958, but by contributing substantial financial assistance to universities, liberal art centers, and national organizations which sponsored and promoted such study-discussion programs. The Great Books Foundation was one of the recipients. Because of the increasing interest in and growth of study-discussion programs in the liberal arts, a need to study the values and effects of these programs developed. Therefore, the Fund for Adult Education authorized a series of research grants for studies of the participants, the leaders, and the educational effectiveness of study-discussion groups. As a result, three evaluation studies were published by the Fund in 1960. One of the studies evaluated the Great Books Program, per se, and members of the Foundation staff cooperated with the researcher in the conduct of the study. Even though these three studies were not conducted under the auspices of the Great Books Foundation, they are included in this investigation because of their purposes and because of the situation as explained in the foregoing paragraphs. Accordingly, in the immediate subsequent sections of this report, the three studies are analyzed after which a summary evaluation is presented.

Kaplan directed the first study which was made in 1956 by members of the faculties of the three test centers: University of California at Los Angeles, The California Institute of Technology at Pasadena, and Whittier College, a small liberal arts college located in Whittier, California.[1]

---

[1] Abbott Kaplan, Study-Discussion in the Liberal Arts

Thus, the groups studied represented those in three cities of different size and character and under different kinds of auspices. The general purpose of the study was to determine the impact of the Liberal Arts Discussion Program on participants and leaders. Primarily, the study endeavored to discover whether intellectual growth, civic participation or continued study had resulted from participation, and whether other social or educational values had accrued from such participation. Answers to several questions were sought. The following were listed:

1. What kind of people enroll in the program?
2. What are the most common reactions of the participants to the program?
3. What relationships exist between certain participant characteristics and certain aspects or effects of the program?
4. How did participants view this experience as compared with more traditional educational experiences they had had?
5. How effective was the leadership?
6. How successful were the discussion groups from the point of view of the leaders?
7. What effect did the experience have on the leaders?
8. How suitable were the materials?
9. To what extent was a subject-matter specialist missed?
10. What is the over-all educational significance and value of such discussion groups?[1]

The study was limited to discussion groups in four

(White Plains, N.Y.: The Fund for Adult Education, 1960).

[1] *Ibid*., pp. 5-6.

programs, namely, "World Affairs Are Your Affairs," "An introduction to the Humanities," "Ways of Mankind," and "World Politics." These programs consisted of 118 discussion groups and involved approximately 2,000 participants during the period of 1954 through the spring of 1956. Each discussion group usually met two hours one evening a week for a period of ten weeks.

Data were obtained from interviews, observations, and from relevant records available at the centers. A random, representative sample of participants and leaders was selected for interviews. This sample included 100 participants from Los Angeles groups; 25 participants each from Pasadena and Whittier groups; and 50 leaders from the three areas. Of the 150 participants, 100 were those who had participated in discussion groups prior to the spring of 1956. They were interviewed once. The remaining 50 were new participants who were enrolled in discussion groups for the first time in the spring of 1956. These 50 participants were interviewed three times: before, during, and immediately following termination of the discussion sessions. Of the 50 leaders, 25 were past leaders and 25 were new leaders. The past leaders were interviewed once, and the new leaders were interviewed twice, before and after the program. In addition, the directors of the test centers were interviewed.

A team of five investigators, including two political scientists, two psychologists, and the research director, conducted the interviews. These investigators also observed four groups for the entire ten sessions, and the first, middle, and last meeting of four additional groups. Each investigator prepared a final report presenting his general observations and estimate of the program. Kaplan stated that "comprehensive questionnaire schedules were used in the interviews."[1] Thus, the study was based primarily on 325 interviews and on observations of 52 group sessions.

Kaplan reported findings descriptively. All data were presented either in frequencies or in percentages or both, and twenty-four tables were interspersed within the text to clarify data at appropriate points in the discussion. The report, a 138-paged book, was divided into six parts. With the exception of Part One, "Introduction," all the report was devoted to a discussion of the findings. The remaining five parts presented findings pertaining to: (1) the participants, (2) views of participatts in the discussion group experience, (3) the leaders, (4) observation of meetings, and (5) summary and conclusions.

---

[1] *Ibid*., p. 7.

Of the seventeen summary statements which were made of the findings, the following are illustrative:

> Participants in the Liberal Arts Discussion Programs for Adults are not typical or representative of the general adult population. They represent a higher-than-average stratum of the population in regard to income, education and occupation. They are active in community activities, read more than average, and read more serious matter.[1]
>
> A majority of the participants believe that members of the groups tend to become more "open-minded" and tolerant of opposing views as a result of their discussion experience. . . . There is little indication that the discussion experience stimulates participants to greater community activity. . . . There is considerable evidence to indicate that participation in the discussion groups stimulates continuing interest and participation in educational and cultural activities.[2]
>
> Despite the weaknesses in materials and leadership, most of the discussion groups are successful in that they provide a stimulating intellectual experience, some growth in knowledge and tolerance of other points of view, as well as in inter-personal relationships. They further stimulate continuing interest and participation in educational and cultural activities.[3]

The conclusions were listed similarly but in greater length. Of the sixteen, the following were included:

> A striking characteristic of participants in the discussion groups is their high economic and educational level. In addition, a high proportion typically engage in educational and cultural activities, more than 50 per

---

[1] *Ibid.*, p. 128. [2] *Ibid.*, p. 129. [3] *Ibid.*, p. 130.

cent having taken classes or courses in the three preceding years, and approximately one-third having been in discussion groups before.[1]

It is evident that the discussion groups have some effect on the reading habits of the participants. More than two-fifths of those interviewed report that they either read more than they had previously or read more serious things. In view of the fact that as a group they already tend to read considerably more, and higher quality reading matter, than the average citizen, the effect of the discussion experience on reading habits may be considered important. . . . How much change occurs in members is difficult to assess. More than half the participants report that they have changed some of their views as a result of the discussions. . . .[2]

There would appear to be little question that many of the participants regard the discussion groups as quite a different kind of experience from their more traditional educational experiences. A sizable number (44 per cent) go so far as to say that it is a more valuable experience. In any event, it has a number of unique aspects for the participants: its informality, lay leadership, and emphasis on discussion. . . .[3]

There is little evidence to indicate that the discussion programs stimulate greater participation in community activities or public affairs.[4]

Finally, it may be said that, although skeptical at the outset of the value of discussion groups led by lay leaders, all the investigators are agreed, despite the many weaknesses revealed in the materials and leadership, that the Liberal Arts Discussion Program for Adults performs a valuable educational function. The value of this contribution can be materially increased by improvement of materials and better selection and training of leaders.[5]

---

[1] *Ibid.*

[2] *Ibid.*, p. 132.

[3] *Ibid.*, p. 133.

[4] *Ibid.*

[5] *Ibid.*, p. 138.

Davis directed the second study in the series for the National Opinion Research Center (NORC) of the University of Chicago in 1957.[1] The Fund for Adult Education commissioned NORC to make the investigation of the Great Books program. The major purpose of the study was to assess the effects of participation in the program in order to secure information to guide future policy in the field of adult education. Davis pointed out that three basic questions guided the research. The three questions and the rationale proposed for each are summarized in the subsequent paragraphs.

1. What are the participants like?--Emphasis was placed on the "social characteristics" of subjects, their stage in the life cycle, their memberships, their institutional affiliations, and their self-conceptions. It was felt that it was not only important to know what the effects were, but also what kinds of people were being affected.

2. What are participants' cultural abilities and interests?--Since it has been commonly thought that Great Books participants were specimens of highly educated Americans, the study examined the participants for clues to

---

[1] James A. Davis, A Study of Participants in the Great Books Program (White Plains, N.Y.: The Fund for Adult Education, 1960).

answers to some general problems that go beyond the scope of the program itself. The literature contains much about "highbrows" and "lowbrows" but little about the cultural life of the "middlebrow," the person who is the Book-of-the-Month reader, the occasional concert-goer, the subscriber to Harper's and The Saturday Review. Thus, since the group of Great Books participants consisted of 1,909 such people, the study sought basic information on their cultural characteristics and orientations, and the patternings thereof.

3. What are the effects of Great Books on the participants?--This problem was approached in two ways. First, the participants were asked what they thought the effects were. Second, comparisons were made of beginning and advanced participants. The researcher pointed out that true effects could be ascertained only from a study which followed the same persons through time. But, since such a study was impossible, he used "cross-sectional" data.

Davis described the procedures employed in the selection of the sample as follows. NORC is organized to take probability samples of the general population of the United States. It has a permanent field staff of trained interviewers in a national sample of counties and standard metropolitan areas, known as "primary sampling units." The counties were so selected that interviews may be weighted and national estimates for a cross-section of the general population may be obtained accurately. These estimates, of course, are subject to random sampling error. Thus, the sample for the study was a probability sample, stratified by year of reading, of the Great Books discussion groups which in November and December of 1957, were meeting in the metropolitan areas and counties outside of the metropolitan areas that comprise NORC's "primary sampling units." The final national sample included 172 groups and 1,909 participants. Since each participant of the sampled groups was asked to respond to a questionnaire, the sample was representative of individuals as well as of groups. However, this procedure resulted in the sample of individuals being heavily clustered. It was recognized that bias was introduced since too many cases were drawn from large cities and too few cases from small towns and rural areas. Davis suggested, however, that in judging the seriousness of this bias, that the following considerations be kept in mind:

1. The Great Books program itself is predominately urban.
2. The budget precluded the employ of trained

> interviewers for a true national sample; hence, it was felt preferable to have national coverage, even with an urban bias, rather than to have perfect sampling of a limited geographical area.

To estimate the practical bias in the sample, selected characteristics of respondents were compared in cities of different size. No consistent differences were found for the following characteristics: (1) age, (2) education, (3) feeling that the program had had a high impact, (4) scores on the test of knowledge of the liberal arts and humanities, (5) self-definition as "intellectual" or "non-intellectual," and (6) interest in community affairs. The following differences were found: (1) Jews are less frequent in the smaller towns than in the large urban areas, (2) republicans are more frequent in smaller towns, and (3) married participants are more frequent in smaller towns. Hence, the data tended to over-estimate the proportion of Jews, Democrats, and non-married in the program.

Data were obtained through a self-administered questionnaire returned by 1,909 participants. NORC interviewers distributed and collected the questionnaires while attending the meetings of 172 groups during December, 1957. The specially developed and pretested questionnaire was a comprehensive one which elicited various types of information.

It contained items to be rated on either a three-, four-, five-, or seven-point scale, and items to be completed, matched, or checked. The questionnaire educed identification-type information, data on social characteristics, opinions on national issues, reports of behavior, and answers to indirect questions. It also contained three tests: (1) to measure knowledge of the liberal arts, (2) to measure musical sophistication, and (3) to measure ability to judge poetry. A copy of the questionnaire occupies twenty-six pages of the Appendices to the report. The questionnaires were coded and punched onto IBM cards for analysis. The researcher stated that the report was based on rather extensive statistical analysis of these data.

Davis noted in the "Introduction" that this published report was a condensed version of the original report. Hence, "a good many of the statistical tables and detailed descriptions of some of the analytical procedures used"[1] were omitted. This report, organized around the three basic questions, treated three parts of the original report. Part A, "The Participants," described the 1,909 sample participants in terms of their salient social characteristics and compared them with other members of the American population. The social characteristics included: education, family situation, occupation, self-conceptions, religious preference, and political party preference. The 1950 Census data were used primarily for comparing subjects with other American people. Essentially, this section of the report presented a sociological description of the participants in terms of the roles they play in their families, in communities, and in the larger society. In Part B, "Self-Assessments," the respondents' own statements pertaining to their reasons for joining and to the effects of participation were examined, summarized, and discussed.

---

[1] Ibid., p. 3.

Part C, "Statistical Assessments," presented results of the comparisons made of the statistical scores of beginning and advanced group members. Participants were divided into three groups: (1) those who had completed less than one year, (2) those who had completed one or two years, and (3) those who had completed three or more years, in Great Books. A statistical score was computed for each respondent, regardless of his year of reading. Thus, new members in the program, even in ninth-year groups, received a score of zero on years of exposure. This part, the longest section of the report, was divided into two chapters. The first chapter examined the survey's data on the relationship between exposure to the program and selected aspects of the participants' intellectual lives. The selected aspects involved three different intellectual areas: (1) knowledge of the traditional liberal arts as revealed by the identification test of thirty-two cartoons, (2) esthetic taste in classical music and in poetry, and (3) patterns of reading. These three were considered together to give a profile of the purely "intellectual" effects of the program. Chapter two considered the program's effects in the areas of basic values, of ideological positions pertaining to religion and politics, and of involvement in local community affairs.

Numerous findings were reported as the researcher discussed and interpreted the data. A summary of the findings was presented at the close of each part of the report or chapter as identified previously. Of the many reported findings, the following, drawn from the summaries, are illustrative:

> In short, the participants are well-educated, high status socially active, youngish adults. . . . They are not, despite their high level of education and later data on their intellectual abilities, so immersed in ideas and culture that these have become the center of their lives. The stereotype of the bookish, ivory

> tower intellectual, does not apply to these people, most of whom share basic middle-class values and few of whom consider themselves intellectuals. The creative intellectual professions are rare among them, and ties to the world of civic organization and political party are strong. . . .[1]
>
> The clearest and strongest effect [of the program] was a definite increase in knowledge of the liberal arts, as measured by a quiz taken by the members. In the area of reading, we found the respondents to be characterized by a consistent pattern of "middle-brow" reading tastes. Our evidence suggests that, in the advanced years, the program does lead to a quantitative increase in hours spent on serious reading, both of Great Books assignments and of self-chosen materials, but that the level of difficulty of the respondents' readings is not related to exposure. Neither of our measures of esthetic taste--musical sophistication and ability to judge poetry--showed any relationship to exposure. . . .[2]
>
> We found no evidence that the members change in their basic values, in the priority they give to such large-scale abstractions as hedonism, contemplation, group affiliation, or tangible activity. At the level of specific ideological positions, however, we think we did spot some directions of change. Continued exposure to Great Books does seem to affect intellectual orientations toward religions and also seems to lead to change in certain political positions. . . . Involvement in the local community shows little change. The members come to the program highly involved, although typically more interested in the larger national or world scene than in the local milieu.[3]

In the concluding chapter, "Summary," the researcher states:

---

[1] *Ibid.*, pp. 25-26.

[2] *Ibid.*, p. 93.

[3] *Ibid.*, p. 109.

> The following paragraphs summarize the major findings of our survey of the Great Books program. In the report itself, these conclusions are developed from rather extensive statistical analysis and are subject to a number of qualifications and specifications which we have left out of the summary.[1]

Thus, conclusions as drawn were essentially a recapitulation of the findings. The following statements were included:

> Great Books attracts people with a wide diversity of motivations, from those who want to learn speed-reading to those who want to solve the world's problem. . . . It is the combination of the social and the intellectual--not one or the other--which appears to be the hallmark of the participants' motivations. . . .[2]
>
> The strongest and clearest effect of continued participation in Great Books is increased knowledge of the liberal arts and humanities. Even when a number of statistical controls are applied, the advanced year participant is considerably more knowledgeable than the first year participant. After three years or more in the program the participant with no college training scores about as well on the knowledge test as the beginning member who has a bachelor's degree. . . .[3]
>
> Any attempt to achieve a final balancing of the books here is beyond the commission of the researcher, for it requires decisions about the relative weight of the high success areas--increased knowledge, changes in ideologies, and participants' subjective satisfactions; the areas of smaller change, such as reading and community activity; and the areas of no effect at all, esthetics and basic values. In addition, even if such a balance could be determined, we would have no idea whether another program or a modified version of the current program would do as well, or better. Such final conclusions will have to be drawn by the reader on the basis of his own system of values and his evaluation of

---

[1] *Ibid.*, p. 112. [2] *Ibid.*, p. 113. [3] *Ibid.*, p. 114.

the practical possibilities in the world of adult education in the liberal arts and humanities.[1]

Hill directed the third study in this series in 1958.[2] The primary purpose of the study was to assess the relative effectiveness of discussion methods, directed by lay discussion leaders, as compared with lecture methods employed by professional, subject-matter specialist. For purposes of the research, effectiveness was defined in terms of the extent to which the objectives of liberal adult education were achieved. Thus, the research problem was that of answering the question, "What is the relative effectiveness of lay-led discussion methods as compared with professionally-led lecture methods in achieving the objectives of liberal adult education?"[3] The two educational methods were compared with respect to the following three general areas of behavioral development or change:

1. The development of mental abilities or skills;
2. Changes in values, interests or attitudes; and
3. Increased knowledge.

---

[1] Ibid., p. 115.

[2] Richard J. Hill, A Comparative Study of Lecture and Discussion Methods (White Plains, N.Y.: The Fund for Adult Education, 1960).

[3] Ibid., p. 5.

The course selected for study was an introductory course in anthropology, "The Ways of Mankind." This course was considered a successful and normal part of the curriculum of the Liberal Arts Adult Education Program of University Extension at the University of California, Los Angeles. It could be treated by both discussion and lecture methods. The discussion series, as usually conducted, incorporated dramatic recordings and a two volume set of readings selected from the professional literature. Discussion groups met for a two hour session weekly for a period of eleven weeks.

The investigator pointed out that the statement of the problem in general terms provided a framework within which the outlines of research could be formulated, and that a more detailed analysis was necessary before making the specific decisions required in the actual execution of the project. Thus, careful specifications were stipulated in the development of the research design; all were met. First, the research was directed by four major goals:

1. . . . this project was designed to provide a reliable description of the relative efficacy of the lecture and discussion methods employed in the Ways of Mankind Series.
2. . . . the project sought to subject to further test certain generalizations and hypotheses derived from previous research in this area.

> 3. . . . the project was designed to explore particular problems in an effort to provide a basis for additional research activity.
> 4. . . . a variety of research techniques were employed in an exploratory effort to evaluate the applicability of these techniques in this type of research situation.[1]

Hill further pointed out that

> on the one hand, the present research is most appropriately viewed as a "pilot study." On the other hand, the project sought to provide reliable information with respect to certain issues currently being faced by those in adult education.[2]

The population included twelve discussion groups, ranging in size from twenty-two to twenty-eight members with a median size of twenty-four, two small lecture classes of twenty-five students in each, and one large lecture group of 233 students. In addition, the following requirements were met:

1. The general subject matter was the same for both the discussion groups and the lecture classes.
2. The subject matter was treated in eleven two-hour sessions.
3. Each meeting of all fifteen groups or classes incorporated a half-hour dramatic recording.
4. The same reading material was used by all participants.

All lecture classes were held at a university, and eleven of the twelve discussion groups met off-campus.

---

[1] *Ibid*., p. 10. [2] *Ibid*., p. 11.

Considerable effort was devoted to developing a number of hypotheses. Some were derived from the social-psychological and educational research literature; others were formulated through the conceptual analysis of the general research problem. In the latter case, the reasoning was mainly speculative, though guided by previous research. Often, a general hypothesis was translated into a set of sub-hypotheses which could be tested directly; at other times, alternative hypotheses were posed. Of the many hypotheses thus formulated, the following are illustrative:

1. Participants who enroll in discussion groups will have had more formal education than those who enroll in lecture classes.
2. No difference will be observed between lecture and discussion participants in terms of the amount of information acquired during the course of the program.
3. During the course of the program, participants in lecture classes will acquire more information than is acquired by participants in the discussion series.
4. As a result of participating in the series, discussion group members will evidence greater attitudinal change than do members of lecture classes.
5. During the course of the program, no change in the attitudes of the participants will be observed, regardless of whether these participants are members of a discussion group or members of a lecture class.
6. Discussion group members will evidence greater

change on the Riecken democracy scale than will lecture participants.

7. As a consequence of social pressure, the more the individual's attitude differs from the norm of his group, the greater the likelihood that the individual will leave the group.

8. The relationship specified in 7 will be more pronounced in discussion groups than in lecture groups.

Data were secured from the study population by three methods: questionnaire, interview, and direct observation. Two questionnaires were developed, the pre- and post-questionnaire. The pre-questionnaire elicited data on the expectations of the participants, selected attitudes, knowledge of anthropological concepts, previous experience in adult education, and socioeconomic characteristics. The post-questionnaire included measures of attitudes, judged effects of the program, evaluations of various aspects of the program, and knowledge of anthropological concepts.

To measure the effects of the program on the attitudes of the participants, the pre-questionnaire included fifty attitude statements, divisible into six attitude scales. These items were randomly arranged, and the participant responded to each statement in terms of a scale ranging from +3 (I agree very much) to -3 (I disagree very much). Thus, a score was computed for each subject on each of the six attitude scales. The attitude scales may be described briefly as follows:

1. The Two Ethnocentrism Scales included a sample of thirteen items employed from the Ethnocentrism Scale by Adorno and others. Seven of the original twelve items were drawn from the minority subscale, and six of the original ten items from patriotism subscale. Items from the two subscales were treated separately.
2. The Authoritarian-Equalitarian Scale consisted of the seven items constructed by R. N. Sanford, one of the original authors of the longer form.
3. The Democracy Scale included ten of the fifteen-item democracy scale developed by Riecken.
4. The Intolerance of Ambiguity Scale included eight of the nine items isolated by Webster, Sanford and Freedman in connection with their studies of authoritarianism.
5. The Attitude toward Adult Education Scale included twelve items. Two forms of this scale was specially developed for this project and pretested.

The Authoritarian-Equalitarian Scale was included only in the pre-questionnaire; all other measures appeared on both questionnaires.

To measure the amount of information acquired during the program, a test calling for the identification of anthropological concepts was included in each of the questionnaires. These tests, specially prepared for the study, were not identical. The pretest consisted of ten matching items, and the posttest included twelve matching items. In order to determine the relative difficulty of the two tests, they were administered to undergraduate students enrolled in Introductory Sociology or General Anthropology during the final week of these courses. Because the mean score attained on the pretest and on the posttest was essentially the same, it was concluded that although the tests were not identical, they were of equal difficulty.

During the first two weeks of the program, semi-structured interviews were conducted with a sample of approximately 20 per cent of the enrolled population. The first interview, averaging one hour and fifteen minutes in length, focused on the respondent's motivation for participation in the series, his past experience in discussion groups, the extent to which he participated in the activities of his community, his organizational affiliations and activities, his reading and recreational activities, and his evaluation of his role in informal group situations. The second interviews, conducted after the conclusion of the program, concentrated on the effects of participation and on the subject's evaluation of the program.

Different procedures were employed in the selection of the interview sample for the three sub-populations, namely, discussion groups, large lecture class, and small lecture classes. The procedures were as follows.

1. Discussion groups.--First, the total discussion population was stratified according to group membership. Secondly, six members were randomly selected from each group. Finally, for each group the six members falling in the sample were assigned randomly to one of three interviewing patterns: (1) two to be interviewed only at the beginning of the program, (2) two to be interviewed only at the end of the program, and (3) two to be interviewed at both times.

2. The Large Lecture Class.--The basic sample of sixty-three participants was selected randomly from those enrolled. These sixty-three subjects were assigned randomly to one of the three interviewing patterns described in preceding paragraph.

3. The Small Lecture Classes.--Only one interviewing pattern was used. Twenty-five participants were selected randomly and interviewed both at the beginning and at the end of the program.

Thus, the total interview pattern was as follows: Twenty-four members of the discussion group and twenty-one members of the large lecture class were interviewed at the beginning of the program. Twenty-four members of the discussion groups and twenty-one members of the large lecture class were interviewed at the end of the program. Twenty-four members of the discussion group, twenty-one members of the large lecture class, and twenty-five members of the small lecture class were interviewed at both the beginning and at the end of the program. The post-interview sample was expanded to include twenty-five participants who were members of two specific discussion groups, one which had maintained a high level of attendance, and one in which attendance had declined to less than 25 per cent of the original group. This special interview sample of twenty-five was selected to provide information which possibly would further the researcher's understanding of factors associated with successful discussion procedures. All interviews were conducted by a team of well-trained investigators who were thoroughly familiar with the project and with the purposes of the research. Post-interviews were completed with 136 subjects.

The third method by which data were obtained was the direct observation of both the lecture and discussion groups. Investigators attended each meeting of the fifteen groups and systematically noted such aspects as attendance, participation, topics discussed, and particular facts of lecturer (or leader) behavior. Also, each of the 165 meeting-observations reports contained the observer's over-all evaluation of the meeting, his judgment as to the degree of interest manifested by participants, a description of the general manner in which the leader or lecturer conducted the session, and a judgment as to the emergence of leaders in discussion groups.

A copy of the interview schedules, forms 1 and 2, and of the pre- and post-questionnaire was included in the Appendices of the report. Of the 576 adults enrolled in the program, 484 completed the pre-questionnaire and 288 completed the post-questionnaire. Comparisons with respect to observed changes in participants were based on the completed pre- and post-questionnaires. The report indicated that data had been subjected to sophisticated analytical treatments.

Various socioeconomic characteristics, experiences, and expectations of participants were reported and compared in terms of the two sub-populations. Primarily, the lecture and the discussion methods were compared in terms of (1) improvement in the ability to identify concepts of anthropology, (2) changes in certain attitudes, and (3) changes in several areas of behavior. The statistical significance of the differences found between the two adult education methods was determined at the .05 and higher levels of confidence. Fifty-two tables presenting the data were interspersed within the text and twenty-four tables were placed in the Appendices.

Essentially, the report was a descriptive interpretation of the findings. It was appropriately organized into eight chapters, five of which were devoted to interpretations of findings as follows: (1) "General Characteristics of the Participants," (2) "Previous Participation in Adult Education Programs and Reasons for Enrolling in This Series," (3) "Attendance," (4) "The Evaluation of the Program by the Participants," and (5) "Effects of Participation in Series." Each of the five chapters contained a concisely-written summary.

Of the many findings reported, the following is a summary of major ones with respect to the illustrative hypotheses stated in a preceding section of this report.[1]

1. No significant differences were found between discussion and lecture participants with respect to educational attainments. Thus, hypothesis 1 was rejected.
2. With respect to the anthropological concepts, the

---

[1] *Supra*, pp. 271-272.

lecture sub-population had a slightly higher average score than did the discussion sub-population at the beginning of the program. At the end of the series, the situation was reversed. Neither of these differences was statistically significant. Thus, hypothesis 2 was accepted, and hypothesis 3 was rejected.

3. With respect to changes in attitudes, the data were not completely clear in terms of hypotheses 4 and 5 as stated. In nine of the ten cases, the observed changes could not be considered statistically reliable. In the tenth case, attitude toward adult education among discussion group participants, a small but significant change occurred, but this change was in an unexpected direction. That is, discussion group subjects were significantly less favorable toward adult education att the end of the series than they were at the beginning. The small changes in the four remaining attitudinal areas, though not statistically significant, were in a consistent direction. Thus, in this study neither method had much effect on the mean attitudes of the groups studied.

4. Both sub-populations became slightly more accepting of democratic group procedures as measured by the Riecken democracy scale. The change for the discussion group participants was greater than that for the lecture group, but the difference was not statistically significant.

5. With respect to hypotheses 7 and 8, the findings indicated that social pressures were operative in both the lecture and discussion situations. However, the data suggested that "attitudinal deviants" are more likely to withdraw from a discussion group than are participants who hold positions at or near the norm for their group. This tendency was not observed for the lecture situation.

The closing chapter of the report, "Conclusion," presented a brief and elucidating interpretation of findings, raised questions, and pointed up problems suggested by the study which should be researched. The following excerpts are illustrative:

> The research as a whole gives little comfort or support to those with strong convictions about the superiority of either the discussion or lecture method of teaching. The over-all impression that one obtains from these results is that the methods investigated are about equally effective. . . .[1]
>
> Possibly the most important consequences of the present study will be in terms of the suggestions that can be derived from it for further research in the adult education area. . . . Liberal adult education courses like those represented by the Great Books Program and the Liberal Arts Programs offered by the University of California do not attract a cross-section of the American public. It is the professional, the educated, and the economically established who enroll in these programs. . . . What of the other segments of the population? Why do the programs that have been studied fail to attract the "blue-collar" worker, the individual with only a grammar-school education, or the economically marginal individual? If one objective of existing liberal adult education programs is to reach as wide a public as possible, a study of program appeal would appear to be in order. . . .[2]
>
> With regard to teaching methodology, it is important to repeat that the findings of the present study must be substantiated before they provide an adequate basis for administrative decision. This research dealt with one course, at one time, in one location. While we believe that our results do have some degree of generality, they should be checked on other programs, at other times and in other places. . . .[3]
>
> Consider our findings with respect to attitudinal change. Only slight changes were observed. If the

---

[1] _Ibid_., p. 84. [2] _Ibid_., p. 85. [3] _Ibid_., p. 86.

> participants continued in adult education programs, would this trend continue? Does participation in liberal adult education programs have a cumulative effect on attitudes? . . . What is required is a longitudinal study, utilizing panel techniques, which follows the same sample of participants through several years of experience.[1]

The three foregoing studies by Kaplan, Davis, and Hill, respectively, are the most comprehensive studies which attempt to evaluate an adult education liberal arts program, in these cases, the small discussion-study groups.[2] The research methodology employed for each may be considered appropriate. However, the reports of these studies are primarily descriptive and are directod more to disseminating information on the findings than they are to describing in detail the research methodology. This makes it more difficult for the writer to analyze these studies accurately and to evaluate them in terms of the specified elements of research.

Even though the three studies were separately conceived and independently conducted, the methodology employed was similar. For each study a team of social scientists was employed to assist with the conduct of the research;

---

[1] Ibid.

[2] Supra: Kaplan, p. 253; Davis, p. 259; Hill, p. 268.

evidently, these teams were highly competent to carry on their respective responsibilities. It was reported that

> Together they [the three studies] represent the first serious attempt to apply the methods of social science research to the evaluation of adult education programs: in this case, programs of reading and discussion in small groups led by non-professional students of the subject matter rather than by experts in it.[1]

The samples for the three studies may be considered adequate. Each of the three researchers selected a random sample which appeared to be representative of his respective population. Kaplan did not indicate the number of leaders drawn from each of the three centers; instead, he reported that the fifty leaders who were selected were from the three areas. However, he stated that his sample was representative. Davis described the procedure employed by NORC to select probability samples of the general population of the United States and recognized that an urban bias was introduced into his national sample. However, to estimate the extent of sampling bias, he compared selected characteristics of respondents in cities of different size and reported that differences were found only for three of the nine characteristics. Furthermore, he pointed out that the

---

[1] C. Scott Fletcher, "Preface," in Study-Discussion in the Liberal Arts by Abbott Kaplan, op. cit., p. iii.

Great Books program itself is predominately urban and that it was preferable to have national coverage with an urban bias than to have perfect sampling of a limited geographical area. Hill used a random sample of approximately 20 per cent of his population for the interviews. Of the 576 participants, 288 completed both the pre- and post-questionnaires. Thus, some bias was introduced into the data since those failing to respond did not do so by chance but because of other operating factors.

Data for the three studies were obtained through the use of one or more of three methods: questionnaire, interviews, and observations. Kaplan reported that personal interviews of two hour duration were held with all members of the sample, and that all sessions of four groups were observed plus additional observations of the first, middle, and final meeting of four other groups. However, Kaplan did not report the kinds of information sought during the interviews or the purpose of the observations. The report contained no schedules for either method. However, the reader, from a careful examination of the research report, is able to identify many, if not all, of the kinds of questions posed during interviews. One may assume that the interviews were semi-structured, and that other questions were directed to a participant depending on his responses and degree of enthusiasm. But, one may suspect that different interviewers may have obtained or interpreted similar responses differently. Also, from the report, a reader perceives some of the kinds of things for which the observers were looking. Kaplan's report presented a seemingly accurate account of information obtained from the interviews and the observations, and it contained an elucidating discussion of particular instances and of reactions and comments by members. Nonetheless, according to the report, an assessment of the quality of the measures cannot be made except to state that only opinions were obtained. Thus, one must conclude that the measure was of poorer quality.

Davis obtained his data by questionnaire which had been pretested with two discussion groups. Elaborate statistical analyses were made of the three tests which were used to assess the effects of the program. The complete questionnaire was included in the report. Thus, it is reasonable to state that the measures employed by Davis were of high quality.

Hill employed all three methods, questionnaire, interview,and direct observations. He described in detail the selection of test items included in the questionnaire, explained that the interviews were semi-structured, and reported at least some of the information contained in the observation reports. Because of the original instruments from which the test items were selected, the measures used by Hill may be considered of high quality. Complete questionnaires and interview schedules were included in the report.

Davis and Hill each applied sophisticated analytical treatments to their data. To assess the effects of the program, Davis was forced to use cross-sectional data obtained from students who had been enrolled for various periods of time. But, careful comparisons were made between beginning and advanced subjects. Hill compared two adult education methods, discussion and lecture. Thus, he tested the effects of group discussion against those of the lecture. In both the Davis and Hill studies, the .05 and higher levels of confidence levels were used to designate significance. Conclusions as drawn were sound and warranted for the studies by Davis and Hill.

Kaplan's study was based on the reported responses of the participants and data were reported descriptively in percentages. If one may assume that each respondent reported accurately, the conclusions may be warranted for this particular group of subjects. However, the study offers no basis for generalizing results to other populations.

These three studies make basic contributions to our knowledge about the participants in the Great Books Programs and of the outcomes of these programs. The work of Davis permits the comparison of Great Books participants with the general population. Their design permits the generalization of the findings to the broad field of liberal cultural programs of which those examined were a carefully drawn sample.

It may be suggested that these studies may well serve as models for the evaluation of local programs in the same and related areas. While the full replication of any one of the studies may be beyond the resources of a local director, portions of any or all may be utilized. Thus local programs, or selected aspects of local programs, may be evaluated and the results related to the findings reported by any of the three authors cited.

CHAPTER IV

RELATIONAL STUDIES BY INSTITUTIONS

The investigation revealed that essentially no evaluation research pertaining to the relational function of adult education has been conducted by the specified institutions. The reader will recall that the relational function of adult education was defined as programs for which objectives are oriented to the development of a deeper understanding of those factors affecting human relationships, and to the improvement of one's adjustment in his personal contacts of life.[1] One unpublished evaluation study of such a program was located. It was conducted under the auspices of University Extension. Thus, the purpose of this chapter is to report the analysis and evaluation of the research report identified which relates to the relational function of adult education.

University Extension

Sonstegard and Rhum investigated the effects of group counseling of dysfunctioning children and of their parents, and the extent to which resulting behavioral changes in the parents and in the children modified the patterns of performance of the children.[1] The study was based on the following hypothesis:

---

[1] Supra, p. 2.

> Academic under-achievement of children may be partially overcome by means of separate group counseling sessions for parents and children designed to effect a change in parents' faulty relationships with their children and the redirection of the children's mistaken goals.[2]

The ten most acute underachieving children were selected from a class of thirty fifth grade children in a laboratory school. A comparison group was formed by matching each member of this experimental group with another child in the same fifth grade class who had not been referred for counseling. The bases for matching were sex and total intelligence score as obtained on the California Test of Mental Maturity, Elementary Form, 1957 edition. The investigators recognized that degree of achievement was not considered as a factor in the matching. Parents of the ten children in the experimental group were invited to form a group for counseling, and eight of the ten sets accepted. This group met one evening each week during the last five months of the 1959-60 school term, and again during the first five months of the 1960-61 school term. A total of thirty counseling sessions were held, each of which was conducted by the junior investigator. The classroom teacher and the current student teacher also participated in the counseling sessions. Group counseling sessions for pupils in the experimental group extended over the same period as that for the group counseling of the parents. This group met thirty minutes one day each week for "Group Discussions for Personal Growth." Neither the pupils in the control group nor their parents participated in counseling sessions.

---

[1] Manford Sonstegard and Gordon Rhum, "Group Counseling of Parents and Their Underachieving Children" (Southern Illinois University, 1963). (Mimeographed.)

[2] *Ibid*., p. 3.

The primary criterion measure to assess the effectiveness of group counseling of parents and of children was the reading achievement of the children. The Iowa Tests of Basic Skills, Forms 1 and 2, were administered before and after treatment, respectively. At the close of the project comparative data were available for six pupils. These data were reported in grade equivalent scores, and three tables presented test results. The investigators found that during the year before counseling the experimental group had a mean per pupil growth of seven months. The corresponding figure for the control group was one year and two months. During the year of counseling the experimental group showed a mean achievement gain of one year and six months; whereas, the mean achievement gain for the control group was just less than one year. Thus, the typical change in the experimental group was +0.9; the corresponding shift for the control group was -0.2.

The written comments and opinions obtained from the parents indicated that the counseling sessions were "generally valuable," and that they had acquired an improved understanding of their children. The teacher observed positive changes in personal conduct and in work habits of the children, and he attributed at least some of the changes to the increased understanding gained by the parents which had led to more effective methods of discipline. The teacher also believed that participation in the counseling sessions had increased his effectiveness in the classroom, and that the counseling sessions were a valuable means of teaching student teachers to understand children.

The investigators pointed up that conclusions drawn from these data should be considered preliminary ones; however, they believed that their findings warranted the tentative suggestion that the method has merit as a means of helping underachievers live more successfully in the intermediate grades. They recommended further investigation.

This study may be viewed as an experimental study employing two groups. As such the research methodology as reported is inadequate. The sample was inadequate in that at the termination of the treatment comparative data were available only for six pupils. Too, reading achievement was reported in grade equivalent scores only, and differences in gains in scores were not tested for significance. The report gave no information on procedures used to obtain parents' written comments and opinions. Neither did the report indicate that any information was obtained from parents of the control group of children. Thus, favorable opinions expressed by parents may not have been a direct result of the counseling sessions. Furthermore, the hypothesis as stated is open to serious questioning since it could not be tested. The report made no reference to it other than to state that the study was based on the hypothesis. Thus, it appears that it was an assumption rather than an hypothesis.

The one measure which was used to ascertain reading achievement may be considered of high quality. However, to attack such an important and perplexing problem, other criterion measures should have been used. Since data were not treated statistically, no confidence levels were specified. The preliminary conclusions as stated may be considered warranted, since the investigators recognized limitations and recommended further investigation. However, conclusions are not warranted from such a study as conducted and reported.

This investigation also may be viewed as a case study of a small group of subjects. As such it served the purpose of an exploratory study. It yielded certain insights which may serve as a basis of more carefully structured work. Future investigators may turn to this study as a source of hypotheses for further study. Its contribution to scholarship if used in this way promises to be much greater than if its results alone are considered.

CHAPTER V

SUMMARY AND CONCLUSIONS

In the foregoing three chapters, the writer has analysed and evaluated the research reports dealing with the evaluation of selected aspects of adult education. The purpose of this chapter is to characterize the present status of evaluation research as research. Because of the manner in which the research reports were selected,[1] it is logical to assume that the forty studies described in this study represent a reasonably good sample of evaluation research in adult education conducted by six institutions.

The analyses and evaluations of the forty research reports reveal that this evaluation research in adult education may be classified in three research categories. These categories are not ordered necessarily in terms of research sophistication; rather, they represent designs which are used frequently in research. The three research categories are:

---

[1] Supra, pp. 11-18.

1. Descriptive studies of single programs;
2. Comparative studies of two or more programs or treatments;
3. Comparative studies of outcomes of the same program on two or more populations.

In terms of these research designs, another possible category is a combination of Numbers 2 and 3; that is, comparative studies of two or more programs or treatments on two or more populations. However, no studies were found in which programs and populations were tested simultaneously.

Descriptive Studies of Single Programs

The descriptive studies of single programs ranged from studies which were essentially survey-type reports to those which were planned analytical investigations of the processes and of outcomes of single programs. Of the some twenty descriptive studies of single programs, one-half were *ex post facto* survey-type reports which were based on minimal data. Data were obtained through the use of questionnaires which had not been tested for validity or reliability. These questionnaires were either mailed to respondents or used during interviews. Thus, data were of undetermined quality, and for the most part, they represented the unsupported opinions and judgments of program participants. All data were reported in frequencies or percentages, and no attempts were made to generalize findings. The investigators indicated that results were a direct influence of the several programs even though no control groups were used. These studies may have been of value to the investigators, particularly in terms of providing needed information for the improvement of the specified programs. Indeed, they were generally so evaluated by their authors. However, they do not contribute to knowledge either in terms of evaluation methodology or of principles generally applicable to adult education.

Six studies in this group represent a higher level of research than do the status-survey reports; yet, statistical treatments were not applied to the data. Because of the way in which data were obtained, these studies may be considered of a higher quality than were the ones limited to status-survey reports. For example, trained interviewers were employed to conduct interviews with respondents in two situations. In three other instances, self-administered questionnaires were reinforced either with observations, interviews with nonparticipating community leaders, tape recordings, or with a test incorporated within the questionnaire. In another study, a test was administered before and after the program, but results were reported descriptively. Essentially, these studies represent the type of descriptive research that is necessary sometimes before testable hypotheses may be formulated. An excellent example is the study directed by Kaplan.[1] His report is a narrative-descriptive type which contains extensive quotations from the respondents, several elucidating brief "case-histories," and percentage tabulations of types of replies received. Many hypotheses based on his survey may be formulated and tested through future research.

Four studies in this group were planned analytical investigations of the processes and of outcomes of a single program. The criterion measure employed was the concept of adoption. Essentially, the primary purpose was to examine one of two different kinds of phenomena which affect the adoption of technological change: sociological factors which correlate with adoption or with nonadoption, and those informational sources which are of most value during the five-stage adoption process. Schedules were used during interviews to obtain data, and all data were subjected to appropriate statistical treatment. Three of these studies tend to be limited in applicability to groups of the type, and to situations of the type studied. One, however, attempted to build and to test a basic body of theory that will account for adoptive behavior.[1] Rogers drew upon basic sociological theory and attempted to identify a minimum number of conceptual variables which together are maximally related to adoption of technological change. His study produced six variables whose multiple correlation with technological change was .409. Rogers' work contributes to knowledge of adoptive behavior in general rather than to knowledge of such behavior in a carefully circumscribed setting.

---

[1] *Supra*, p. 253.

Comparative Studies of Two or More Programs or Treatments

A little less than one-half of the reported research involved those studies in which two or more programs or treatments were compared. For the most part the basic design, per se, for these studies meet acceptable criteria for research designs. However, one or more research procedures were violated in some ten of the studies.

The purpose(s) as stated for a few of the studies were general. It was evident that objectives of the program were not clearly defined. The questionnaires used to obtain data elicited unverified opinions and judgments; consequently, the data derived from such instruments are of questionable reliability and validity. In a few studies the sample was either quite small or nonrepresentative of the population studied. In others, the samples to be compared were unmatched. These errors introduce bias into the data. The criterion measures which were used in some four or five studies were invalid in terms of the objectives. In most of the same studies, no base lines were established for participants in terms of the phenomena being measured; thus, the true effects of treatments could not be assessed. For several of the studies, inappropriate statistical treatments or no statistical treatments were applied to the data. Despite these weaknesses as research, many of these studies do afford information from which testable hypotheses may be formulated.

---

[1] *Supra*, p. 35.

The remaining studies in this group, with minor exceptions, were neatly designed and clearly reported investigations. Samples were adequate; they were either random samples of a well defined universe, or they included all subjects in two established groups who could be matched on important variables. The criterion measures were valid particularly in terms of the stated objectives of the research. Thus, data were of high quality. A base line was established, and treatment affects were assessed in terms of subjects' movement from this base line. All data were subjected to analytical treatments which were sensitive and appropriate for determining differences in change. The confidence levels of .05 and higher were accepted for indicating significance. Data and results from statistical analyses were reported appropriately. Findings were clearly reported and conclusions, as drawn, were sound and warranted in terms of reported findings. In fact, these studies may be duplicated by other researchers interested in evaluating a similar adult education program or in comparing two or more adult education methods. The studies conducted by Rogers and Havens, Marsh and Hyman, Welch, Costin and Johnston, Anikeeff, Schwartz, Hadlock, and Hill are highly creditable pieces of research which deserve emulation.[1]

All of the studies in this category contribute to practice except one in which a theoretical evaluative technique was tested.[1] Schwartz developed and tested a criterion measurement based on the logic of semantic differentials. His study primarily contributes to knowledge since it opened up a theory that may be tested in other situations.

---

[1] Supra: Rogers and Havens, p. 88; Marsh and Hyman, p. 96; Welch, p. 107; Costin and Johnston, p. 146; Anikeeff, p. 148; Schwartz, p. 151; Hadlock, p. 212; Hill, p. 268.

Comparative Studies of Outcomes of Same Program on Two or More Populations

Two of the forty studies represent research which compares the affects of the same program on two or more populations. These studies were well-designed and meet acceptable standards of research. Random samples were carefully selected and every effort was made to determine the extent to which they were representative of the populations studied. Criterion measures were valid and they were particularly sensitive to the selected phenomena which were measured. Thus, data were of high quality. Also, they were subjected to quite sophisticated analytical treatments. Characteristics of subjects were carefully examined in terms of differences in and among populations studied. The change in subjects was determined by the use of sensitive statistical treatments, and the .05 and higher levels of confidence were used to designate significance of differences. Conclusions were sound and warranted in terms of reported findings.

---

[1] Supra, p. 151.

The study directed by Davis is the most comprehensive evaluation study of an adult education program.[1] His national sample, stratified by years of participation, afforded cross-sectional data which may be considered to represent three populations. The questionnaire he used consisted of some 300 items, many of which were answerable in graded options of from two to seven choices. These data were subjected to a variety of detailed, exhaustive, and sophisticated statistical techniques. Future researchers may turn to Davis' study for guidance and apply the same criterion measures to compare the same subjects over an extended period of time or to compare two or more treatments on different populations.

In the other study an hypothesis was formulated and tested.[2] Brandner hypothesized that one population which had been successful in adopting a recommended agricultural innovation would adopt another recommended innovation which was similar more rapidly than would a second population which had had no experience with the previously recommended innovation. To test his theory, he selected two populations for both of which the adoption of the newly recommended practice would be an economic asset. The first population had previously adopted a similar innovation, but the second had not previously adopted such similar innovation. From the literature on adoption he identified a large number of factors that had been shown by other investigators to be positively related to adoption. Of these, three factors significantly indicated that the first population would adopt the new practice and nine that the second population would be the adopters. Thus on the basis of the number of factors favoring each, the odds were only three to nine that the population having previous familiarity would lead in adoption. The results of his experimental treatment showed, however, that the members of the population having prior experience in adopting a similar innovation actually adopted the new practice in the ratio of four to every one of the population having no previous similar experience. In actuality the economic advantage of adopting would have been greater for the non-adopting group than for the group that did adopt. His operating hypothesis was supported by a confidence level of .001. Thus, Brandner's study contributes to knowledge rather than to practice since the underlying theory may explain adoptive behavior in other situations.

---

[1] Supra, p. 259. [2] Supra, p. 27.

Conclusions

The status of evaluation research may be summarized as follows:

1. The majority of studies are highly localized and whatever value they possess is limited to the program studied; i.e., the findings obtained from such studies are not generalizable.
2. Many studies are severely limited by weakness of design and of sampling, inadequacy of the data, and inadequacy of the analysis of the data that were obtained.
3. Some studies, approximately 20 per cent of those reviewed, yielded findings that are generalizable.
4. Some studies, approximately 20 per cent of those reviewed, may well serve as models directly, or with minor remediation, for evaluative research in adult education.
5. A few studies, not more than 15 per cent of those reviewed, made fundamental contributions to the general theory of adult education or to the evaluation of adult education programs.

It seems clear that evaluation studies which are of a purely local character should be appraised in terms of their operational values to the institution making them and not as research in any fundamental sense of the term. Researchers in adult education evaluation research should pay particular attention to the elements of good design, to sampling techniques, to instrument development, and to appropriate statistical treatments of their data. Even more fundamentally, evaluations should be conducted within the framework of a coherent body of theory, and the results should lead to the further extension and refinement of the theory within which the researcher is working.

The critical examination of the evaluation research covered in this study underscored the fact that adult education is inherently complex. As a totality, it draws upon a number of basic disciplines among which are sociology, cultural anthropology, social psychology, learning theory, and one or another of any number of content areas, and educational methodology. An adequate evaluative research study necessarily involves competent scholarships in those areas which are germaine to the program under study, plus scholarly competency in research design, measurement, and statistical analysis. The weaknesses noted throughout this study are probably attributable to the fact that the investigator did not have at his disposal the full array of competencies needed. It seems unrealistic to expect that many individual adult educators would possess all needed competencies. It therefore seems that evaluation research in this field should employ the team approach, associating scholars in the pertinent fields in the total effort. It also seems that financial resources far exceeding those apparently available to most investigators will be required for the basic and comprehensive research which is needed.

BIBLIOGRAPHY

I. *General References*

Adult Education Association of the U. S. A. *Adult Education*, Vols. 1 - 15, 1950-1965.

_______. *Adult Leadership*, Vols. 1 - 13, 1952-1965.

_______. "Research and Investigations in Adult Education," *Adult Education*, Vols. 5 (1955), pp. 114-127, 240-246; 6 (1956), pp. 234-243; 7 (1957), pp. 196-207; 8 (1958), pp. 195-207; 9 (1959), pp. 198-210; 10 (1960), pp. 195-218; 11 (1961), pp. 195-222; 12 (1962), pp. 195-216; 13 (1963), pp. 195-216; 14 (1964), pp. 195-225; and 15 (1965), pp. 195-234.

Aker, George F. *Adult Education Procedures, Method and Techniques*. A Classified and Annotated Bibliography, 1953-1963. New York: University College, Syracuse University, 1965.

American Educational Research Association. "Adult Education," *Review of Educational Research*, Vols. 23 (1953); 29 (1959); and 35 (1965).

Brunner, Edmund deS., *et al*. *An Overview of Adult Education Research*. Chicago: Adult Education Association of the U. S. A., 1959.

Klineberg, Otto. "The Problem of Evaluation," *International Social Science Bulletin*, Vol. 7 (1955), pp. 346-352.

Knowles, Malcolm S. (ed.). *Handbook of Adult Education in the United States*. Chicago: Adult Education Association of the U. S. A., 1960.

Kreitlow, Burton W. "Research in Adult Education," in Handbook of Adult Education in the United States. Edited by Malcolm S. Knowles. Chicago: Adult Education Association of the U. S. A., 1960.

Lionberger, Herbert F. Adoption of New Ideas and Practices. Ames, Iowa: The Iowa State University Press, 1960.

Little, Lawrence C. A Bibliography of Doctoral Dissertations on Adults and Adult Education. Revised ed. Pittsburgh: University of Pittsburgh Press, 1963.

Miller, Harry L. "Philosophy and Issues," Review of Educational Research, Vol. 24 (1959), pp. 230-238.

Rogers, Everett M. Diffusion of Innovations. New York: The Free Press of Glencoe, 1964.

University of Wisconsin. "Research in Cooperative Extension Work." A series of manuscripts, one prepared annually since 1957. Department of Agricultural and Extension Education, College of Agriculture, 1957-1964. (Mimeographed.)

U. S. Department of Health, Education, and Welfare. Summaries of Studies in Agricultural Education. Prepared by the Research Committee of the Agricultural Education Section, American Vocational Association. Vocational Division Bulletin Nos. 248 (1952); 251 (1953); 253 (1954); 256 (1955); 263 (1956); 265 (1957); 272 (1958); 255 (1959); 282 (1960); 291 (1961); and 300 (1962).

U. S. Federal Extension Service. Review of Extension Research. Extension Service Circular Nos. 470 (July, 1950); 471 (Jan., 1951); 474 (July, 1951); 480 (March, 1952); 486 (Jan., 1953); 493 (Jan., 1954); 497 (Feb., 1955); 506 (May, 1956); 511 (June, 1957); 518 (July, 1958); 521 (July, 1959); 532 (July, 1960); 534 (July, 1961); 541 (Sept., 1962); and 544 (Oct., 1963).

Wilder, David E. "Problems of Evaluation Research," in An Overview of Adult Education Research by Edmund deS. Brunner et al. Chicago: Adult Education Association of the U. S. A., 1959.

II. Evaluation Research

A. Cooperative Extension

1. Behavioral Change

Alexander, Frank D., and Longest, James W. *Evaluation of the Farm Management Phase of the Farm and Home Management Program in New York State*. Study No. 1. Ithaca: New York State University Extension Service, Cornell University, 1962.

References: Report Nos. 1-12. Each of the twelve mimeographed bulletins is a separate but comprehensive report on some aspect of the evaluation study of the five-year experimental program in Farm and Home Management. (Available: Office of Extension Studies, Roberts Hall, Cornell University).

Bass, Bruce Carter. "The Relationship of Systematic Instruction in Vocational Agriculture for Young Farmers in Virginia to Certain Factors of Establishment." Unpublished Ed.D. dissertation, Pennsylvania State University, 1955.

Reference: *Agricultural Education Magazine*, Vol. 28 (November, 1955), pp. 109; 115.

Boggs, Carl D. "An Evaluation of the Institutional On-Farm Training Program for Veterans in Oconee County." Unpublished M.S. thesis, Clemson Agricultural College, 1951.

Bond, T. H. "An Evaluative Study of a District Extension Program in Marketing." Unpublished M.S. thesis, Michigan State University, 1961.

Breeden, Laurie Leroy. "A Study of the Institutional On-Farm Training Program in St. Helena Parish." Unpublished M.S. thesis, Louisiana State University, 1951.

Busbice, Walter E. "Changed or Improved Practices that Have Grown Out of the Gaithersburg Young Adult Farmer Program." Unpublished M.S. special problem, Department of Agricultural Education, University of Maryland, 1956.

Button, D. J. "An Audience Inventory of Illinois Farmers: A Guide to Improve Communication of Dairy Information." Unpublished M.S. thesis, University of Illinois, 1957.

Chandler, Carrol W. "Evaluation of On-Farm Training at Carter High School." Unpublished M.S. problem, Department of Agricultural Education, University of Tennessee, 1951.

Clem, Margaret Louise. "An Evaluation of the Present Status of the Craft Phase of the Tennessee Agricultural Extension Service Program in Upper East Tennessee To Be Used as a Basis for Future Program Development." Unpublished M.S. thesis, University of Tennessee, 1961.

Cochran, Lacy E. "A Study of the Progress of a Group of Veterans Enrolled in Institutional On-Farm Training in Hardy County, West Virginia." Unpublished M.S. thesis, West Virginia University, 1954.

Condon, L. M. "Extension Programs in Estate Management." Unpublished M.Ed. thesis, Colorado State University, 1959.

Cook, Lowell Nathan. "Progress Made by Thirty Veterans Enrolled in the Institutional On-Farm Training Program in Clay County, West Virginia University, 1952.

Craig, Leland Wiseman. "The Effect of the Institutional On-Farm Training Program toward Establishment in Farming." Unpublished M.S. thesis, Mississippi State University, 1956.

Dorries, William Lyle. "A Study of the Effectiveness of the Veterans' Farm Training Program in Fannin, Hunt, and Lamar Counties in Texas." Unpublished M.S. thesis, Texas A. and M. College, 1952.

Felton, Robert A. "The Institutional On-Farm Training Program in Tucker County, West Virginia." Unpublished M.S. thesis, West Virginia University, 1951.

Fessenden, J. G., and Rohrer, W. C. *A Study of an Urban Home Economics Extension Program, Baltimore, Maryland*. Miscellaneous Publication 159. College Park: University of Maryland Agricultural Extension Service, 1957.

Fields, Addison J. "Changes and Improvements Made by Madison County, Alabama, Negro Veterans on their Farms and in Their Farming Programs While Enrolled in the Institutional On-Farm Training Program." Unpublished M.S. thesis, Ohio State University, 1952.

Fish, David William. "A Study of 157 Pennsylvania Institutional On-Farm Trainees in Their Progressive Establishment in Farming; Practices, and Outcomes of the Training Program, as Related to Future Programs in Agricultural Education." Unpublished M.Ed. problem, Department of Agricultural Education, Cornell University, 1951.

Foster, George Samuel. "An Evaluation of Supervision of Volunteer Leaders in Tennessee 4-H Club Work." Unpublished Ed.D. dissertation, George Washington University, 1962.

Frank, Harry Elmer, Jr. "The Development and Evaluation of an Adult Farmer Educational Program in the Tryon, Oklahoma Community." Unpublished M.S. thesis, Oklahoma State University, 1957.

Furbay, Walter Merrill. "An Evaluation of the Program of Adult Education in Vocational Agriculture at Clay-Genoa School." Unpublished M.S. thesis, Ohio State University, 1956.

Griffin, M. *An Evaluation of the Farm and Home Development Program in Massachusetts, 1955-1956*. Amherst: Massachusetts University Agricultural Extension Service, 1956.

Guiler, Gilbert Samuel. "An Evaluation of the Young and Adult Farmer Program at Canal Winchester, 1950-51." Unpublished M.A. thesis, Ohio State University, 1951.

Hardee, J. G. "Evaluation of an Educational Program with Part-time Farm Families. Transylvania County, North Carolina." North Carolina State College Agricultural Extension Service, 1961. (Mimeographed.)

Reference: North Carolina Extension Evaluation Studies No. 2, 1957.

Hargrave, Harvey M. "Determining the Progress and Changes in Farm Organization as They Relate to the Livestock Program among Members of the East Arcadia, North Carolina, Adult Farmers' Class." Unpublished special problem, Department of Agricultural Education, Pennsylvania State University, 1955.

Hendrickson, V. C. "An Appraisal of the Rural Resource Development Program in Price County, Wisconsin." Unpublished M.S. thesis, University of Wisconsin, 1960.

Holder, Nadine D. "Evaluation of the Home Demonstration Club Program in Meeting the Interests and Needs of Urban Homemakers of Washington County, Oklahoma." Unpublished M.S. thesis, Oklahoma State University, 1959.

Holik, J. S., and Moore, C. R. "Balanced Farming Evaluation." University of Missouri Agricultural Extension Service, 1957. (Mimeograhped.)

Holmes, M. S. "Some Indications of Level of Knowledge and Opinions of Michigan Homemakers about Food Buying." Unpublished M.S. thesis, Michigan State University, 1958.

Huffman, Joseph Carlyle. "A Study of the Effectiveness of the Institutional On-Farm Training Program in Virginia." Unpublished M.S. thesis, Cornell University, 1951.

Hutchinson, Harvey E. "Some Factors Determining Effectiveness of the Institutional On-Farm Training in Tangipohoa Parish." Unpublished M.S. thesis, Louisiana State University, 1950.

Jennings, Sybil. "Studying Home Demonstration Work with Rural Homemakers in Winn Parish." Unpublished M.S. thesis, Louisiana State University and Agricultural and Mechanical College, 1954.

Johnson, Donald E., and Wilkening, Eugene A. _Five Years of Farm and Home Development in Wisconsin: A Comparison of Participating and Control Families_. Research Bulletin 228. Madison: University of Wisconsin Agricultural Experiment Station, 1961.

References: Preliminary Reports Nos. 1 (1956), 2 (1957), and 3 (1958).

Lacy, Micajah Pennington. "The Effects of Involvement on the Participants in Cooperative Extension Program Planning in Waupaca County, Wisconsin." Unpublished Ph.D. dissertation, University of Wisconsin, 1961.

Lintner, Julius H. "A Study of the Effectiveness of the Institutional On-Farm Training Program in Ohio with Implications for Future Programs in Adult Education." Unpublished Ph.D. dissertation, Ohio State University, 1952.

McKenzie, George S. "An Evaluative Study of the Veterans' Farm Training Program of the Seneca High School." Unpublished M.S. thesis, Clemson Agricultural College, 1951.

Marsh, C. Paul. *An Evaluation of the Farm and Home Development Approach to Agricultural Extension Work in North Carolina*. Extension Evaluation Studies No. 3. Raleigh: North Carolina State College Agricultural Extension Service, 1962.

Reference: North Carolina Extension Evaluation Studies No. 1, 1957.

Matott, H. W. "A Study of the Farm Management Phase of the Farm and Home Management Program in Chenango County." Unpublished M.S. thesis, Cornell University, 1958."

Matthews, J. L., and Whitham, G. E. *Cooperative Agricultural Extension Work in Windham County, Connecticut*. Storrs: Connecticut Agricultural College Extension Service, 1953.

Mayeaux, John Samuel. "Some Accomplishments of Institutional On-Farm Training toward Establishment in Farming." Unpublished M.S. thesis, Louisiana State University, 1951.

Minter, E. M. *Money Management Workshops for Young Homemakers . . . A Study in Three Indiana Counties*. Research Bulletin 731. Lafayette: Indiana Agricultural Experiment Station, 1961.

Nickell, E. "A Study to Determine Homemakers Gain in Knowledge as a Result of Attending Six Meetings on Adjusting Homemaking to Today's Living." Unpublished special problem report for Course EPD 899, Michigan State University, 1961.

Nielsen, James, and Bittner, R. F. *Farm Practice Adoption in Michigan*. Technical Bulletin 263. East Lansing: Michigan Agricultural Experiment Station, 1958.

Nielson, James, and Crosswhite, William. *The Michigan Township Extension Experiment: Changes in Agricultural Production, Efficiency and Earnings*. Technical Bulletin 274. East Lansing: Michigan State University Agricultural Experiment Station, 1959.

References: Technical Bulletins 266 (1958); 284 (1961); 287 (1962), and two subsequent reports.

Norby, Oscar W. "An Appraisal of Long-Time Cooperative Extension Program Planning in Waupaca County, Wisconsin." Unpublished Ph.D. dissertation, University of Wisconsin, 1961.

Parker, R. H. "A Study of the Effectiveness of the Institutional On-Farm Training Program in the Mars Hill Community." Unpublished M.S. thesis, Mississippi State University, 1953.

Phillips, Robert Allen, Jr. "A Study to Evaluate and Improve the Adult Farmer Program in the Morris Chapel Community." Unpublished M.S. thesis, University of Tennessee, 1954.

Preston, J. C. "An Evaluation of the Farm and Home Management Program in Genesee County, New York." *Unpublished M.S. thesis, University of Wisconsin*, 1959.

Raaf, V. A. "Ten Years of Balanced Farming in Usage County, Missouri, from 1947 to 1957." Unpublished M.S. special report, University of Missouri, 1957.

Ramsey, Ralph J. *Evaluation of the Kentucky Farm and Home Development Program, 1951*. RS-1. Lexington: University of Kentucky Agricultural Extension Service, 1952.

Rohrer, W. C. *The Extension Service in Cecil County, Maryland*. Miscellaneous Publication 216. College Park: Maryland Agricultural Experiment Station, 1954.

St. John, George William. "A Study of the Changes Made in Dairy Management Practices by Clark County Farmers on Spencer Silt Loam Soil Type from 1949 to 1957." Unpublished M.S. thesis, University of Wisconsin, 1959.

Samendinger, H. "The Cornell Farm Study Course Program." Unpublished M.S. thesis, Cornell University, 1955.

Silverman, L. J., and Bailey, W. C. *Trends in the Adoption of Recommended Farm Practices, Alcorn County, Mississippi, 1954-1957*. Bulletin 617. State College: Mississippi Agricultural Experiment Station, 1961.

Slocum, W. L., and Brough, O. L., Jr. *Family and Farm Changes Associated with Farm and Home Planning in Washington*: Bulletin 633. Pullman: Washington Agricultural Experiment Station, 1962.

References: Preliminary Report (1956); Bulletin 584 (1958); Circular 403 (1962).

Soults, M. W. "Summary Report of Changes Made by Couples Participating in the Farm and Home Development Program." MS-714. Iowa State University Agricultural Extension Service, 1958. (Mimeographed.)

References: One or two interim reports of this study are available.

Spraggs, Paul Forest. "Changes in the Conditions and Situations of Veterans Enrolled in Institutional On-Farm Training, Halifax Training School, Halifax, Virginia." Unpublished M.S. thesis, Virginia State College, 1955.

Strand, Forrest Oliver. "An Evaluation of the Institutional On-Farm Training Program in Michigan." Unpublished M.A. thesis, Michigan State University, 1953.

Swayze, James Enoch, Sr. "A Study of Institutional On-Farm Training for Negro Veterans at Zebulon, North Carolina." Unpublished M.S. thesis, Ohio State University, 1952.

Taylor, Louis C. "An Evaluation of the Adult Education Program in Vocational Agriculture in Tennessee during 1952-53." Unpublished M.S. thesis, University of Tennessee, 1954.

Winterfeldt, E. P. "Evaluation of the Home Demonstration and 4-H Girls Programs in Pittsburg County, Oklahoma, Based on Opinions of Home Demonstration Club Officers." Unpublished M.S. thesis, University of Wisconsin, 1959.

2. Effectiveness of Method(s)

Ascheman, R. E. "The Use of Result Demonstrations by Minnesota County Agricultural Agents." Unpublished Plan B Paper (M.S. in Agricultural Education), University of Minnesota, St. Paul, 1959.

Axinn, George H. *Farm Audience in Delaware*. Circular 4. Newark: Delaware Agricultural Experiment Station, 1952.

Bertrand, Alvin L., and Bates, Frederick. *Television in Rural Louisiana*. Bulletin 518. Baton Rouge: Louisiana Agricultural Experiment Station, 1958.

Borton, R. E. "Consumer Use of Mass Media for Food Information." Unpublished M.S. thesis, Michigan State University, 1957.

Reference: *Quarterly Bulletin of the Michigan Agricultural Experiment Station*, Vol. 41, No. 1 (August, 1958), pp. 24-36.

Burgess, C. *Consumer Marketing Information Survey*. Berkeley: California University Agricultural Extension Service, 1957.

Carpenter, Earle S. *"Gardener's Almanac" Television Series for Home Gardeners.* Auburn: Massachusetts University Agricultural Extension Service, 1958.

Carroll, C. B. "A Normative Survey of Viewers of the County Agent's Daily Television Program in Caddo Parish, Louisiana." Unpublished M.S. thesis, Louisiana State University, 1959.

Clyatt, G. "A Survey of the Effectiveness of the Extension Consumer Marketing Information Program When Transmitted through Mass Media." Unpublished M.S. thesis, University of Houston, 1962.

Cooley, Esther. "The Effectiveness of Home Furnishing Clinics." Louisiana University Agriculture Extension Service, 1950. (Processed)

________. *Informing Consumers in Lake Charles: An Evaluation*. Publication 1233. Baton Rouge: Louisiana University Agricultural Extension Service, 1958.

Coolican, P. M. "A Study of the Effectiveness of Teaching by Television Versus Teaching by the Use of an Extension Bulletin." Unpublished M.S. thesis, Michigan State University, 1960.

Coombs, K. L. "Agricultural Extension Meetings as a Medium for Farm Information." Unpublished M.S. thesis, University of Maryland, 1954.

Donahoo, Alvin William. "Differenetial Attitudes of Farmers toward the Grain Marketing System." Unpublished Ph.D. dissertation, University of Minnesota, 1954. (*Dis. Absts*., Vol. 15, p. 371)

Erickson, S., and Olien, C. Evaluation of TV Clothing Series "What Shall I Wear?" St. Paul: Minnesota University Agricultural Extension Service, 1961.

Eschler, Richard E., Dell, J. C., Jr., and Alexander, F. D. *Evaluation Study of the TV Dairy Cattle Feeding School*. Special Report 6. Ithaca: Cornell University Agricultural Extension Service, 1962.

Federal Extension Service. Better Informed Consumer . . . A Report on Evaluation Study of the Extension Service Consumer Marketing Information Program at Louisville. Circular 502. Washington, D. C., 1955.

References: Federal Extension Service Circulars 499 and 501.

Gibbs, M. A. *Indications of Homemakers Motivations and Knowledge about Food Buying*. East Lansing: Michigan State University Extension Service, 1959.

Guarino, Russell. "Evaluation of Monthly Dairy Discussion Group Meetings as an Agricultural Extension Teaching Method in LaFayette Parish." Unpublished M.S. thesis, Louisiana State University, 1961.

Holloway, M. E. *An Evaluation of the Maryland Extension Fair Exchange TV Program of September 17, 1960*. Baltimore City Extension Service, 1961.

Iowa Agricultural Extension Service. *Make a Dress--TV!* Ames, 1953.

La Frenier, B. M. "A Study of the Effectiveness of Television in Teaching Clothing Construction." Unpublished M.S. thesis, University of Tennessee, 1960.

Lionberger, Herbert F. *Television Viewing in Rural Boone County with Special Reference to Agricultural Shows*. Research Bulletin 702. Columbia: University of Missouri Agricultural Experiment Station, 1958.

Marsh, C. Paul, and Hyman, Theodore M. Programmed Learning in Extension Training. A Comparison of the Effectiveness of Programmed Self-Instruction and Workshop Instruction. Technical Bulletin 161. Raleigh: North Carolina Agricultural Experiment Station, 1964.

Merrill, I. R. "Impact of the 'County Agent' Television Film Series," Michigan Agricultural Experiment Station Quarterly Bulletin, Vol. 40 (May, 1958), pp. 758-769.

Musgrove, G. M. "Homemakers' Evaluation of Extension Teaching Methods." Unpublished professional paper for course, Ex. 171: Extension Practicums, Colorado State University, 1958.

Pollock, J., and Meloche, G. The Effectiveness of Television in Teaching Tailoring a Coat. Madison: University of Wisconsin Home Economics Extension Service, 1954.

Porter, W. F., et al. Food Buying: Knowledge--Concerns--Practices. An Evaluation of the Marketing Information for Consumers Program in the Wheeling-Steubenville Area. Bulletin 456. Morgantown: West Virginia Agricultural Experiment Station, 1961.

Porter, Ward S., and Wilson, Lorene S. Evaluation of St. Joseph Consumer Marketing Program. Extension Circular 784. Columbia: University of Missouri Agricultural Extension Service, 1962.

Reuter, Ava Marie. "An Evaluation of the Leader Training System of Home Demonstration Work in Wisconsin as Viewed by Home Agents." Unpublished M.S. thesis, University of Wisconsin, 1959.

Robert, James Joseph. "An Evaluation of Selected Teaching Methods in Getting a New Farm Practice Adopted in St. Landry Parish, 1960." Unpublished M.S. thesis, Louisiana State University, 1961.

Ronken, C. P. *A Study of the Consumer Marketing Information Program, Jackson, Mississippi*. Publication 434. State College: Mississippi State University Agricultural Extension Service, 1962.

Rogers, Everett M., and Havens, A. Eugene. *The Impact of Demonstrations on Farmers' Attitudes toward Fertilizer*. Research Bulletin 896. Wooster: Ohio Agricultural Experiment Station, 1961.

Sarbaugh, L. E. "Measuring the Effectiveness of an Information Campaign on Soil Testing in Bond County, Illinois." University of Illinois, College of Agriculture, 1959. (Typewritten.)

Schaeffer, A. J. "A Study of the Comparative Effectiveness of Three Communication Channels Used by a Cooperative Extension Agent in Teaching Homemakers." Unpublished M.S. thesis, Michigan State University, 1960.

Schwartz, L. M. "A Comparison of Traditional and Mechanical Aid Training Methods in an Extension Home Economics Situation." Unpublished M.S. thesis, Michigan State University, 1961.

Simonds, L. A. "Sources of Information Used and Kinds of Help Desired in Planning and Purchasing Food for Selected Families in Loraine County, Ohio." Unpublished M.S. thesis, Ohio State University, 1958.

Tait, Eldon B. "A Comparison of Self-Instruction and a Workshop as Methods of Training Newly Hired County Extension Personnel in the Subject of Radio Broadcasting." Unpublished M.S. thesis, Cornell University, 1962.

Tait, Eldon B., *et al*. *Radio Experiment in Reaching Working Women*. Extension Studies 1. University Park: Pennsylvania State University Extension Service, 1958.

Welch, John M. "An Evaluation of Three Adult Education Methods for Disseminating Trade Information to Missouri Restaurant Operators." Unpublished Ph.D. dissertation, Florida State University, 1961.

References: *Adult Education*. Vol. 12 (Summer, 1962), pp. 231-237; *The Cornell Hotel and Restaurant Administration Quarterly*, May, 1962.

Welker, Walter. "The Cooperative Extension Television Audience in a Pennsylvania Rural Area." Unpublished M.S. thesis, Pennsylvania State University, 1962.

Reference: *The Evaluator*, No. 18, October, 1962. (University Park: Pennsylvania State University Agricultural Extension Service)

Willsey, Frank Richard, Jr. "An Experimental Study of an Adult Learning Situation Involving Three Levels of Training in the Group Discussion Process." Unpublished Ed.D. dissertation, Indiana University, 1962. (*Dis. Absts*., Vol. 23, p. 2407)

Wilson, D. D., *et al*. *A Study of the Food for Health Project*, San Diego County, 1954. Berkeley: California University Agricultural Extension Service, 1954.

Wilson, Meredith C., and Moe, Edward O. *Effectiveness of Television in Teaching Sewing Practices*. Federal Extension Service Circular 466. Washington, D. C., 1951.

3. The Adoption-Diffusion Process

Abell, Helen C. "The Differential Adoption of Homemaking Practices in Four Rural Areas of New York State." Unpublished Ph.D. dissertation, Cornell University, 1951.

Adams, Dewey Allen. "Factors Associated with the Use of Approved Farm Practices in the Beaver Dam Community of Cumberland County, North Carolina." Unpublished problem option (Master's in Agricultural Education), North Carolina State College, 1958.

Alexander, F. D. "Studying the Decision-Making Process," in *The Research Clinic on Decision Making*. Pullman: State College of Washington (August, 1958), pp. 21-35.

Anderson, Marvin A. "Factors Affecting Acceptance and Use of Fertilizer on Iowa Farms." Unpublished Ph.D. dissertation, Iowa State University, 1955.

Reference: Special Report 16. Ames: Iowa Agricultural Experiment Station, 1956.

Armstrong, J. B. "County Agent Activities and the Adoption of Soil-Building Practices." Unpublished M.S. thesis, University of Kentucky, 1959.

Reference: Kentucky Agricultural Experiment Station Progress Report 130, 1963.

Averill, Thomas B. "The Relationship between the Educational Activities of Adults and the Differential Adoption of Farm Practices." Unpublished Ph.D. dissertation, University of Chicago, 1962.

Bailey, W. C., *et al*. *Community Structure and Farmer Education*. Sociology and Rural Life Series 8. State College: Mississippi Agricultural Experiment Station, 1957.

Bauder, Ward. *Influences on Acceptance of Fertilizer Practices in Piatt County*. Bulletin 679. Champaign: Illinois Agricultural Experiment Station, 1961.

Beal, George M. "Information Sources in the Decision-Making Process," in *The Research Clinic on Decision Making*. Pullman: State College of Washington (August, 1958), pp. 36-51.

Beal, George M., and Rogers, Everett M. *The Adoption of Two Farm Practices in a Central Iowa Community*. Special Report 26. Ames: Iowa Agricultural Experiment Station, 1960.

________. "Informational Sources in the Adoption of New Fabrics," *Journal of Home Economics*, Vol. 49 (1957), pp. 630-634.

________. "The Scientist as a Referent in the Communication of New Technology," *Public Opinion Quarterly*, Vol. 22 (1958-59), pp. 555-563.

Beal, George M., Rogers, Everett M., and Bohlen, Joe M. "Validity of the Concept of Stages in the Adoption Process," *Rural Sociology*, Vol. 22 (1957), pp. 166-168.

Bonser, H. J. "Rural Community Clubs in Tennessee--Their Bearing on Adoption of Farm and Home Practices." Unpublished Ph.D. dissertation, Pennsylvania State University, 1957.

References: Bulletins 286 and 287. Knoxville: Tennessee Agricultural Experiment Station, 1958.

Brandner, Lowell. "Evaluation for Congruence as a Factor in Accelerated Adoption of an Agricultural Innovation." Unpublished Ph.D. dissertation, University of Wisconsin, 1960.

Reference: *Rural Sociology*, Vol. 24 (1959), pp. 381-383.

Burleson, G. L. *Studying Extension Work with Farmers and Farm Homemakers in Madison Parish, Louisiana*. 2 Vols. Baton Rouge: Louisiana University Agricultural Extension Service, 1952.

Burleson, G. L. *Studying Extension Work with Farmers and Farm Homemakers in Red River Parish, Louisiana*. 2 Vols. Baton Rouge: Louisiana University Agricultural Extension Service, 1950.

________. *Studying Extension Work with Farmers and Farm Homemakers in Washington Parish, Louisiana*. 2 Vols. Baton Rouge: Louisiana University Agricultural Extension Service, 1950.

Chaparro, Alvaro. "Role Expectation and Adoption of New Farm Practices." Unpublished Ph.D. dissertation, Pennsylvania State University, 1955.

Coleman, A. Lee. "Differential Contact with Extentension Work in a New York Rural Community," *Rural Sociology*, Vol. 16 (1951), pp. 207-216.

Coleman, A. Lee, and Marsh, C. Paul. "Differential Communication among Farmers in a Kentucky County," *Rural Sociology*, Vol. 20 (1955), pp. 93-101.

Copp, James H. *Personal and Social Factors Associated with the Adoption of Recommended Farm Practices among Cattlemen*. Technical Bulletin 83. Manhattan: Kansas Agricultural Experiment Station, 1956.

________. "Toward Generalization in Farm Practice Research," *Rural Sociology*, Vol. 23 (1958), pp. 103-111.

Copp, James H., Sill, Maurice L., and Brown, Emory J. "The Function of Information Sources in the Farm Practice Adoption Process," *Rural Sociology*, Vol. 23 (1958), pp. 146-156.

Coughenour, C. Milton. "The Functioning of Farmers' Characteristics in Relation to Contact with Media and Practice Adoption," *Rural Sociology*, Vol. 25 (1960), pp. 283-297.

Coughenour, C. Milton, and Patel, N. B. *Trends in Use of Recommended Farm Practices and Farm Information Sources in Twelve Kentucky Neighborhoods*. Progress Report 111. Lexington: Kentucky Agricultural Experiment Station, 1962.

Cummings, Gordon J. "The Differential Adoption of Recommended Farm Practices among Dairymen in New York Community." Unpublished M.S. thesis, Cornell University, 1950.

Dean, Alfred. "Some Factors Related to 'Rationality' in Decision-Making among Farm Operators." Unpublished M.A. thesis, University of North Carolina, 1958.

Reference: *Rural Sociology*, Vol. 23 (1958), pp. 121-135.

Duncan, James A. "The Relationship of Selected Cultural Characteristics to the Acceptance of Educational Programs and Practices among Certain Rural Neighborhoods in Wisconsin." Unpublished Ph.D. dissertation, University of Wisconsin, 1953.

References: *Rural Sociology*, Vol. 19 (1954), pp. 349-357; Research Bulletin 525. Madison: Agricultural Experiment Station, 1956.

Estep, Allen J. "The Relation of Agricultural Education to Adoption of Farm Practices by Young Farm Operators." Unpublished M.S. thesis, University of Wisconsin, 1958.

Reference: Technical Bulletin 214. Madison: Wisconsin Agricultural Experiment Station, 1959.

Fliegel, Frederick C. "A Multiple Correlation Analysis of Factors Associated with Adoption of Farm Practices." Unpublished Ph.D. dissertation, University of Wisconsin, 1955.

Reference: *Rural Sociology*, Vol. 21 (1956), pp. 284-292.

Fliegel, Frederick C. "Farm Income and the Adoption of Farm Practices," *Rural Sociology*, Vol. 22 (1957), pp. 159-162.

Fosen, Robert Harlan. "Social Solidarity and Differential Adoption of a Recommended Agricultural Practice." Unpublished Ph.D. dissertation, Cornell University, 1956.

________. "Structural and Social Psychological Factors Affecting Differential Acceptance of Recommended Agricultural Practices: A Study in the Application of Reference Group Theory." Unpublished M.S. thesis, Cornell University, 1956.

Felder, Allie Clay. "The Acceptance of Recommended Peanut Production and Marketing Practices in Nansemond County, Virginia, Relative to Family Organization, Family Values and Related Social and Economic Status Factors." Unpublished Ph.D. dissertation, Ohio State University, 1954.

Gray, John H. "Sources of Farming Information Used by Hopkins County Kentucky Farmers." Unpublished M.S. thesis, University of Wisconsin, 1956.

Gross, Neal N., and Taves, Marvin J. "Characteristics Associated with Acceptance of Recommended Farm Practices," *Rural Sociology*, Vol. 17 (1952), pp. 321-327.

Hall, D. M. *A Survey of Land Use Practices and Communication Channels Regarding Soil Information*. Urbana: Illinois University Agricultural Extension Service, 1957.

Harris, Ruth. "Certain Socio-Economic Factors and Value Orientations as Related to the Adoption of Home Practices." Unpublished M.S. thesis, University of Wisconsin, 1956.

Havens, A. Eugene. "Social Psychological Factors Associated with the Differential Adoption of New Technologies by Milk Producers." Unpublished Ph.D. dissertation, Ohio State University, 1962.

Havens, A. Eugene, and Rogers, Everett M. "Adoption of Hybrid Corn: Profitability and the Interaction Effect," *Rural Sociology*, Vol. 26 (1961), pp. 409-414.

Hess, C. V., and Miller, L. F. *Some Personal, Economic and Sociological Factors Influencing Dairymen's Actions and Successes*. Bulletin 577. State College: Pennsylvania Agricultural Experiment Station, 1954.

Hobbs, Daryl J. "Factors Related to the Use of Agricultural Chemicals on Iowa Farms." Unpublished M.S. thesis, Iowa State University, 1960.

Hoffer, Charles R., and Stangland, Dale. *Farmers' Reactions to New Practices*. Technical Bulletin 264. East Lansing: Michigan Agricultural Experiment Station, 1958.

Reference: *Rural Sociology*, Vol. 23 (1958), pp. 112-120.

Kimball, William J. "The Relationship between Personal Values and the Adoption of Recommended Farm and Home Practices." Unpublished Ph.D. dissertation, University of Chicago, 1960.

Kivlin, Joseph E. "Characteristics of Farm Practices Associated with Rate of Adoption." Unpublished Ph.D. dissertation, Pennsylvania State University, 1960.

References: *Social Forces*, Vol. 40 (1962), pp. 364-370; Research Bulletin 691. University Park: Pennsylvania Agricultural Experiment Station, 1962.

Klietsch, Ronald. "Decision-Making in Dairy Farming: A Sociological Analysis." Unpublished Ph.D. dissertation, University of Minnesota, 1961.

Lackey, Alvin S. "The Consistency of Sociological Variables in Predicting the Adoption of Farm Practices." Unpublished Ph.D. dissertation, Cornell University, 1958.

Leuthold, Frank O. "Demonstrators and the Diffusion of Fertilizer Practices." Unpublished M.S. thesis, Ohio State University, 1960.

Reference: Research Bulletin 908. Wooster: Ohio Agricultural Experiment Station, 1962.

Lionberger, Herbert F. "Community Prestige and the Choice of Sources of Farm Information," *Public Opinion Quarterly*, Vol. 23 (1959), pp. 110-118.

________. *Information Seeking Habits and Characteristics of Farm Operators, Based on a Study Conducted in Northeast Missouri*. Research Bulletin 581. Columbia: University of Missouri Agricultural Experiment Station, 1955.

________. *Legitimation of Decisions to Adopt Farm Practices and Purchase Farm Supplies in Two Farm Communities: Ozark and Prairie*. Research Bulletin 826. Columbia: University of Missouri Agricultural Extension Service, 1963.

________. "Reception and Use of Farm and Home Information by Low-Income Farmers in Selected Areas of Missouri." Unpublished Ph.D. dissertation, University of Missouri, 1950.

Reference: Research Bulletin 472. Columbia: Missouri Agricultural Experiment Station, 1951.

________. "Some Characteristics of Farm Operators as Sources of Farm Information in a Missouri Community," *Rural Sociology*, Vol. 18 (1953), pp. 327-338.

Lionberger, Herbert F., and Coughenour, C. Milton. *Social Structure and Diffusion of Farm Information*. Research Bulletin 631. Columbia: Missouri Agricultural Experiment Station, 1957.

Marsh, C. Paul, and Coleman, A. Lee. *Communication and the Adoption of Recommended Farm Practices*. Progress Report 22. Lexington: Kentucky Agricultural Experiment Station, 1954.

________. "Farmers' Practice-Adoption Rates in Relation to Adoption Rates of Leaders," *Rural Sociology*, Vol. 19 (1954), pp. 180-183.

________. "Group Influences and Agricultural Innovations," *American Journal of Sociology*, Vol. 61 (1956), pp. 588-594.

________. "The Relation of Farmer Characteristics to Adoption of Recommended Farm Practices," *Rural Sociology*, Vol. 20 (1955), pp. 289-296.

________. "The Relation of Kinship, Exchanging Work, and Visiting to the Adoption of Recommended Farm Practices," *Rural Sociology*, Vol. 19 (1954), pp. 291-293.

________. "The Relation of Neighborhood of Residence to Adoption of Recommended Farm Practices," *Rural Sociology*, Vol. 19 (1954), pp. 385-389.

Maquin, C. J. "Factors Associated with the Adoption of Selected Dairy Farming Practices, Avoyelles Parish, 1956." Unpublished M.S. thesis, Louisiana State University, 1957.

Olson, K. S. "The Relation of Selected Farmers' 4-H Experience to Their Adoption of Improved Farm Practices." Unpublished Ph.D. thesis, University of Wisconsin, 1959.

Pedersen, Harold A. "Cultural Differences in the Acceptance of Recommended Practices," *Rural Sociology*, Vol. 16 (1951), pp. 37-49.

Poirot, J. A. "Use of Available Sources of Homemaking Information by Rural Homemakers in Iowa." Unpublished M.S. thesis, Iowa State University, 1957.

Ramsey, Ralph J. "A Study of Changes in Value Orientations in an Adult Education Program." Unpublished Ph.D. dissertation, University of Chicago, 1958.

Rogers, Everett M. "The Adoption Period," *Rural Sociology*, Vol. 26 (1961), pp. 77-82.

________. "Categorizing the Adopters of Agricultural Practices," *Rural Sociology*, Vol. 23 (1958), pp. 345-354.

________. *Characteristics of Agricultural Innovators and Other Adopter Categories*. Research Bulletin 882. Wooster: Ohio Agricultural Experiment Station, 1961.

________. "A Conceptual Variable Analysis of Technological Change." Unpublished Ph.D. dissertation, Iowa State University, 1957.

Reference: *Rural Sociology*, Vol. 23 (1958), pp. 136-145.

________. "Personality Correlates of the Adoption of Technological Practices," *Rural Sociology*, Vol. 22 (1957), pp. 267-268.

Rogers, Everett M., and Beal, George M. "The Importance of Personal Influence in the Adoption of Technological Changes," *Social Forces*, Vol. 36 (1958), pp. 329-335.

________. *Reference Group Influence in the Adoption of Agricultural Technology*. Ames: Department of Economics and Sociology, Iowa State University, 1958.

Rogers, Everett M., and Burdge, Rabel J. *Community Norms, Opinion, Leadership, and Innovativeness among Truck Growers*. Research Bulletin. Wooster: Ohio Agricultural Experiment Station, 1962.

________. *Muck Vegetable Growers: Diffusion of Innovations among Specialized Farmers*. Research Circular 94. Wooster: Ohio Agricultural Experiment Station, 1961.

Rogers, Everett M., and Capener, Harold R. *The County Extension Agent and His Constituents*. Bulletin 858. Wooster: Ohio Agricultural Experiment Station, 1960.

Rogers, Everett M., and Havens, A. Eugene. *Extension Contact of Ohio Farm Housewives*. Research Bulletin 890. Wooster: Ohio Agricultural Experiment Station, 1961.

Rogers, Everett M., and Pitzer, R. L. *The Adoption of Irrigation by Ohio Farmers*. Research Bulletin 851. Wooster: Ohio Agricultural Experiment Station, 1960.

Rogers, E. M., and Yost, M. D. *Communication Behavior of County Extension Agents*. Research Bulletin 850. Wooster: Ohio Agricultural Experiment Station, 1960.

Sheppard, David. "Neighborhood Norms and the Adoption of Farm Practices," *Rural Sociology*, Vol. 25 (1960), pp. 356-358.

Sill, M. L. "Personal, Situational, and Communicational Factors Associated with the Farm Practice Adoption Process." Unpublished Ph.D. dissertation, Pennsylvania State University, 1958.

Sizer, Leonard M., and Porter, Ward F. *The Relation of Knowledge to Adoption of Recommended Practices*. Bulletin 446. Morgantown: West Virginia Agricultural Experiment Station, 1960.

Spaulding, Irving A. *Farm Operators' Time-Space Orientations and the Adoption of Recommended Practices*. Bulletin 330. Kingston: Rhode Island Agricultural Experiment Station, 1955.

Spencer, George Elwood. "Value-Orientations and the Adoption of Farm Practices." Unpublished Ph.D. dissertation, Cornell University, 1958. (*Dis. Abstr*., Vol. 18, pp. 1884-1885)

Reference: *Rural Sociology*, Vol. 24 (1959), pp. 35-47.

Straus, Murray A. "Family Role Differentiation and Technological Change in Farming," *Rural Sociology*, Vol. 25 (1960), pp. 219-228.

Tucker, Carlos F. "Prediction of Rate of Adoption from Characteristics of Farm Innovations." Unpublished M.S. thesis, Ohio State University, 1961.

van den Ban, Anne Willem. "Locality Group Differences in the Adoption of New Farm Practices," *Rural Sociology*, Vol. 25 (1960), pp. 308-320.

Wilkening, Eugene A. *Acceptance of Improved Farm Practices in Three Coastal Plain Counties*. Technical Bulletin 98. Raleigh: North Carolina Agricultural Experiment Station, 1952.

________. *Adoption of Improved Farm Practices as Related to Family Factors*. Research Bulletin 183. Madison: Wisconsin Agricultural Experiment Station, 1953.

________. "Changes in Farming as Related to Familism, Family Decision-Making and Family Integration," *American Sociological Review*, Vol. 19 (1954), pp. 37-39.

________. "Informal Leaders and Innovators in Farm Practices," *Rural Sociology*, Vol. 17 (1952), pp. 272-275.

________. "Joint Decision-Making in Farm Families as a Function of Status and Role," *American Sociological Review*, Vol. 23 (1958), pp. 187-192.

________. "Roles of Communicating Agents in Technological Change in Agriculture," *Social Forces*, Vol. 34 (1956), pp. 361-367.

________. "Social Isolation and Response of Farmers to Agricultural Programs," *American Sociological Review*, Vol. 16 (1951), pp. 836-837.

Wilkening, Eugene A., and Johnson, Donald. "A Case Study in Decision-Making among a Farm Owner Sample in Wisconsin," in *The Research Clinic on Decision Making*. Pullman: State College of Washington (August, 1958), pp. 1-20.

________. *Goals in Farm Decision-Making as Related to Practice Adoption*. Bulletin 225. Madison: Wisconsin Agricultural Experiment Station, 1961.

Young, James N. "The Influence of Neighborhood Norms on the Diffusion of Recommended Farm Practices." Unpublished Ph.D. dissertation, University of Kentucky, 1959.

Reference: *Rural Sociology*, Vol. 24 (1959), pp. 372-380.

Young, James N., and Marsh, C. Paul. *The Adoption of Recommended Farm Practices and Sources of Farmer Information*. Progress Report 40. Lexington: Kentucky Agricultural Experiment Station, 1956.

B. *University Extension*

Anikeeff, Alexis M. "Scholastic Achievement of Extension and Regular College Students," *Journal of Applied Psychology*, Vol. 38 (1954), pp. 171-173.

Costin, Frank, and Johnston, Robert L. "A Study of Achievement in an Introductory Psychology Course," Adult Education, Vol. 12 (1962), pp. 120-126.

Hadlock, Alton Parker. "A Study of the Development of Critical Thinking through Discussion Groups." Unpublished Ed.D. dissertation, University of California (Los Angeles), 1958.

McCormick, Fred C. "A Comparative Study of Engineering Extension and Resident Programs," Adult Education, Vol. 11 (1961), pp. 106-109.

Murphy, Thomas F. "The University's Role in Management Development." Unpublished M.B.A. thesis, University of California (Berkeley), 1959.

Schwartz, Michael. "Adult Education to Industry: A Program Evaluation." Unpublished M.A. thesis, University of Illinois, 1959.

Sonstegard, Manford, and Rhum, Gordon. "Group Counseling of Parents and Their Underachieving Children," Southern Illinois University, 1963. (Mimeographed)

C. Evening College

Gibson, James Edward. "A Study of Selected Factors and Attitudes Comparing Regular Day Students with Continuing Education Students." Unpublished Ed.D. dissertation, University of Arizona, 1962.

Manley, Charles Arnold. "An Evaluation of the Extended Day Program at San Fernando Valley State College." Unpublished Ed.D. dissertation, Washington State University, 1961.

D. Junior or Community College

Luchsinger, Leland Beyer. "Criteria for Evaluation of Adult Education Programs in Junior Colleges and Their Application to Kilgore College." Unpublished Ed.D. dissertation, University of Texas, 1959.

E. Public School Adult Education

Burkey, Robert Lyle. "Student Responses to an Evening Session Public School Adult Education Art Program." Unpublished Ed.D. Type C Project, Teachers College, Columbia University, 1961.

Cheskie, Sophie Veronica. "An Appraisal of the Adult Education Program of the Highland Park Public Schools, Highland Park, Michigan." Unpublished Ed.D. dissertation, Wayne State University, 1957.

Matherne, Allen James. "An Analysis and Evaluation of an Experimental Program Designed to Increase the Level of Economic Understanding among Adults." Unpublished Ph.D. dissertation, Michigan State University, 1960.

Shrock, John G. "Functions of Methodology in Adult Education Ceramics Courses." Unpublished Ed.D. dissertation, Stanford University, 1953.

Timken, Joseph E. "A Study to Determine the Adequacy of Three Selected Oklahoma Public School Adult Education Programs." Unpublished Ph.D. dissertation, Oklahoma Agricultural and Mechanical University, 1952.

F. Great Books

Davis, James A. A Study of Participants in the Great Books Program. White Plains, New York: The Fund for Adult Education, 1960.

Hill, Richard J. A Comparative Study of Lecture and Discussion Methods. White Plains, New York: The Fund for Adult Education, 1960.

Kaplan, Abbott. Study Discussion in the Liberal Arts. White Plains, New York: The Fund for Adult Education, 1960.

E N D. DEPT. OF HEALTH EDUCATION & WELFARE U.S. OFFICE OF EDUCATION E R I C.
DATE FILMED 6. 26. 67